Essays on Human Capital Development

John F. Creamer

Submitted for the degree of Doctor of Philosophy

Heriot-Watt University

School of Social Science

Department of Accounting, Economics and Finance

November 2017

Abstract

This thesis comprises three essays which examine how skills gaps form, persist and can be closed in young Latin American children. Skills are the foundation of an individual's human capital, and since the early 2000s have been broadly defined as cognitive (things that are taught) or non-cognitive (things that are learned through experience). The theoretical underpinning of the development of such skills is the production function for skills introduced by Todd and Wolpin (2003, 2007). The setting of Latin America is a relevant context to undertake this research, as it is one of the most unequal regions in the world. The first two studies make use of longitudinal Young Lives data from Peru, while the last uses government data from Honduras.

In the first paper, the analysis exploits the longitudinal aspect of Young Lives to examine when nutritional investment is most productive in the first five years of life. The study looks to improve on the existing literature by using two new instrumental variable methods: weak identification robust confidence intervals, and instruments generated from assumptions about heteroskedasticity. The second paper uses siblings data to estimate effects of parental and peer relationships on early adolescent children's self-esteem and pride. It then decomposes socioeconomic and locality gaps to determine how policy can potentially equalise outcomes in each group. The last paper evaluates the impact of a randomised cash transfer in Honduras on early childhood development outcomes. Specifically, the study examines if a transfer conditional on school attendance promotes intra-household spillovers from the beneficiary siblings to younger siblings under the age of 5. The estimates show modest improvements in early childhood development for these younger siblings. The body of evidence suggests that these improvements are real spillover effects and not just a result of an income effect for beneficiary households. Overall, the thesis provides evidence which supports the theoretical model on the importance of early intervention, while noting the necessity of using sound methodologies to deal with empirical issues in estimating such effects.

Acknowledgements

I am fortunate to have been helped along this journey by so many people. My two supervisors, Cath and Mark, have been instrumental in this process, giving me the freedom to work on my own ideas and the support and feedback to turn these ideas into worthwhile output. Thanks especially to Cath for setting up opportunities to work, present, and being a tremendous resource throughout my MSc Dissertation and my PhD. Thanks to the lecturers and staff at Heriot-Watt who took time to listen to presentations and give feedback and advice. Thanks (in no particular order) to Ryuta, Jan, Fran, Achim, Antonio, Kyle, Wei, Rong, Marco and all who joined the 12pm lunch crew for the respite from the work at HW. I am appreciative of Florencia and the rest of the staff in SPH for making the opportunity happen, and welcoming me with open arms (even if my Spanish left something to be desired).

My friends here in Edinburgh and back in the States have been the source of great relief, experiences, and hilarity. Edinburghers, be it playing music (Sulkas what up!), pub nights, watching (proper) football, or travelling, you have provided laughs and care for someone who's a long way from home. It has meant so much. Much thanks to the Haries, the finest running club in all of Scotland (yes, I went there), for the banter, the Meadows miles, and this new distance running kick. Dr. Ethan Matz (you beat me!), thanks for the medical advice! Justin, I appreciate the visits, the football, and everything else. Hugo, there is a reason why I travelled to Sunderland of all places to get away! Andrew, you saved my PhD! Cameron, Paul, Mackle, Bryan, Brian, Peter, and more (I won't bore you with a full list, you know who you are), your contributions have not been forgotten.

Finally, Jonas and Wizzy; thanks for listening, giving me a title, and putting up with me, my (sometimes) ridiculous training, and my overall nonsense. I would have driven myself crazy without you both. To Mom and Dad, to Daniel and Tonya, and the rest of my family; you will never know how much I appreciated the calls, the FaceTimes, the venting sessions, the visits, the games of FIFA, and the unyielding support of me and my work. I am incredibly grateful, and look forward to closing the distance.

DECLARATION STATEMENT

(Research Thesis Submission Form should be placed here)

Contents

[1]The data for Chapters 2 and 3 come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India (Andhra Pradesh and Telangana), Peru and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development (DFID), with co-funding from 2010 to 2014 by the Netherlands Ministry of Foreign Affairs, and from 2014 to 2015 by Irish Aid. The views expressed here are those of the author and not those of Young Lives, the University of Oxford, DFID or other funders. STATA code available at www.github.com/jcreamer15.

List of Figures

List of Tables

Glossary

ASQ3- Ages and Stages Questionnaire 3

AR- Anderson-Rubin Weak Instrument Robust Test

BDH- Bono de Desarollo Humano

CCT- Conditional Cash Transfer

DFID- United Kingdom Department for International Development

ECD- Early Childhood Development

EGRA- Early Grade Reading Assessment

IV- Instrumental Variable

K- Kleibergen Statistic

OLS- Ordinary Least Squares

PPVT-Peabody Pictoral Vocabulary Test

PRAF I/II- Programa de Asignación Familar I/II

RCT- Randomised Control Trial

RIF- Recentered Influence Function

RPS- Red Protección Social

SD- Standard Deviation

SES- Socioeconomic Status

UK- United Kingdom

US- United States

USD- United States Dollar

1 Introduction[2]

"We cannot always build the future for our youth, but we can build our youth for the future." –Franklin Delano Roosevelt, 1940.

This thesis examines three current issues in the early childhood and human capital development literature, focusing specifically on children's skills development. The topics of the essays cover three specific aspects of child development, yet share a common objective to increase the knowledge of channels of development in the early life to inform policy. The first two chapters focus on two distinct channels of skills development; malnutrition and the home environment. The last chapter examines whether a randomised cash transfer conditional on school attendance spurs intra-household spillovers, offering a contribution to the impact evaluation literature beyond average treatment effects for the treated household member.

The main theoretical framework that is used throughout this thesis is the production function for skills (Todd and Wolpin 2003, 2007). The production function framework is attractive for this research because it treats skills as something that can be developed through investment in nutrition, education and a solid home environment. Cunha and Heckman (2007) advance this method by discussing the importance of investment in the earliest stages of life. In their perspective, skills gaps that are formed early in life are more difficult to close as skills become entrenched. Skills are classified into two different types; *cognitive* skills, which are skills that can be taught, and *non-cognitive* (or soft skills), which cannot be explicitly taught, but are learned through experience and observations of others. These skills have been proven to be important for labour outcomes and personal behaviour in adulthood. Adults with higher stocks of these skills are more likely to have jobs and less likely to be in jail, use drugs, and participate in other risky behaviours.

[2]The data for Chapters 2 and 3 come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India (Andhra Pradesh and Telangana), Peru and Vietnam (www.younglives.org.uk). Young Lives is funded by UK aid from the Department for International Development (DFID), with co-funding from 2010 to 2014 by the Netherlands Ministry of Foreign Affairs, and from 2014 to 2015 by Irish Aid. The views expressed here are those of the author and not those of Young Lives, the University of Oxford, DFID or other funders. STATA code available at www.github.com/jcreamer15.

This thesis uses the above framework to build a greater understanding of the importance of the timing of nutritional investment for cognitive skill development, the role parents and the environment has on children's later childhood non-cognitive development, and how policy can spur development through income transfers and intra-household spillovers. The papers all make methodological advances to the existing literature, illuminating new paths of research. The first chapter makes two important contributions to the literature researching the impacts of early childhood malnutrition on cognitive skill development. First, it promotes the use of weak-identification robust methods when examining results to ensure that inference is a strong as possible. Second, it applies a new instrumental variable (IV) method (Lewbel 2012) using instruments generated from assumptions about heteroskedasticity which can be used when exogenous instruments are unavailable in the data. This is a significant contribution to the literature because it highlights weaknesses in existing methodologies, while also paving a way forward when experimental or quasi-experimental methods cannot be used. The chapter concludes that these new methods provide robustness checks which may lead to better inference in this area of research.

The second chapter exploits data that have been collected on two siblings per household to examine how gaps that emerge in indices of self-esteem and pride behave by socioeconomic status and locality. Additionally, it determines the role parent-child and peer-child relationships have on the development of these two psychosocial competencies which are important for individual's labour market success. These robust estimates of the determinants of self-esteem and pride add to the existing literature by showing the importance of relationships when considering the determinants of self-esteem and pride in early adolescents. The last chapter uses a randomised cash transfer, Bono 10000, as its source of plausible identification. The primary contribution of the evaluation is the identification of indirect early childhood development effects stemming from beneficiary children attaining more education. The effect is seen through an increasing programme effect for heterogeneous interactions based on household size and beneficiary child characteristics. In terms of methodology, the study is one of the first in its field to account for multiple comparisons. Testing a range of different hypotheses increases the likelihood of obtaining a false positive result. We use Romano and Wolf (2005)'s stepwise multiple hypothesis testing method to recalculate p-values controlling for the family-wise error rate, leading to more robust estimates.

The three studies are set in Latin America, a region plagued by high levels of inequality. Poor rural areas have less access to schools and health services, and are more exposed to exogenous shocks in terms of the environment and economy, emphasising the need for intervention to give households the opportunity to participate in the national and global economy. The first two studies use the Young Lives data from Peru, a longitudinal dataset which follows children from the age of 6 to 18 months to 11 and 12 years old (the latest publicly available round of the survey). Young Lives visits households (who were randomly selected at baseline) triennially, collecting information on household and individual characteristics and, important for this research, information on the cognitive and non-cognitive skills development of children. Additionally, the study asks questions of the index children's siblings, providing an opportunity to use a sibling difference estimator to remove household fixed effects. In the Peruvian case, the sample is pro-poor, but representative on the national level except for the upper 5% of the income distribution. The final study uses survey data from Honduras collected by the government in conjunction with the randomised cash transfer that is being evaluated, Bono 10000.

1.1 Outline of the three chapters

The first study, *Identifying the Role of Nutrition in Developing Cognitive Skill in Young Peruvian Children,* examines when nutritional investments are most productive in the first years of life. It aims to model causal effects of early childhood and preschool nutrition on mid-childhood cognitive ability in Peru. Building on the theoretical framework of critical and sensitive periods of human capital development, the empirical specification obtains the total effects of these periods of nutrition. OLS results show a lower bound estimate of the effects of early childhood nutrition consistent with the past literature. The total effect is decomposed into direct and indirect effects, highlighting the importance of nutrition at different stages in the early childhood. Causal effects are difficult to obtain when studying nutrition at two periods of time. Therefore, two additional instrumental variable methods are applied for the first time in this literature. First, weak identification robust methods are used to test the validity of the coefficients in the standard model, illustrating possible causal parameter values based on prior information. Then, instruments are generated from assumptions about heteroskedasticity in the data to create a source of exogenous variation. These estimates are imprecise. The results support the OLS estimates as being lower bounds on the effect of early nutrition on cognitive ability, but that no causal relationship can be confidently stated because of the imprecise estimates. The result provides a cautionary note for the previous literature which has not undertaken robust inference, and suggest a methodological path forward for the future research.

The second study, *The importance of family, friends and location on the development of human capital in mid-childhood and early adolescence,* asks how relationships with parents and peers relate to socioeconomic and locality gaps in two domains of children's psychosocial competencies, self-esteem and pride. A cohort of Peruvian children and their siblings between the ages of 6 and 12 is used for the analysis. Non-parametric analysis maps the differences by wealth and locality in children's pride, self-esteem, and their relationship with peers and parents. Gaps are prevalent, but diminish as children age, in all measures outside of peer relationships. Notably, parent-child relationships in rural households are worse than those in urban settings, illustrating one of the consequences of living in rural areas. Parametric analysis estimates wealth gradients and determinants of self-esteem and pride. The results show that relationships are hugely important, explaining between 50 to 80 percent of the variance in children's self-esteem and pride together. While there are mean gaps, there is no evidence of wealth gradients in this sample for self-esteem and pride. The unique sibling's aspect of the sample is exploited in a siblings difference model to improve identification, showing that the initial estimates are robust to household fixed effects. Decomposition results show that mean gaps in self-esteem and pride by wealth and locality can be closed by improving the parent-child relationship, with improvements of 21% in poor households and 80% for rural households. These estimates suggest that the previous literature which do not measure relationships provide lower bound estimates of the effect of parents on their children's development. Policies which can improve these relationships, especially the parent-child relationship, are important for giving poorer children the essential human capital to overcome poverty in the future.

The final study, *Child development in a large cash transfer: Exploring sibling effects in a Honduran RCT,* explores the effects of a randomly assigned conditional cash transfer in Honduras (Bono 10000) on early childhood development. We find significant impacts on cognitive development (as measured by the language subscale of ASQ) in children 0-60 months old, with an effect size of 0.17 SD overall. The observed improvements in outcomes do not seem to be explained by the increase in income alone. As much as the programme appears to have slightly changed some behaviours that benefited children (higher expenditures on nutritious food, such as proteins and vegetables, healthcare and schooling, and increased maternal empowerment and self-esteem), we show that these are not the channels behind the positive impacts of the cash transfer. Sibling spillover effects appear as an important channel: young children with at least one eligible older sibling in the household have significantly larger effects of the transfer on their language development of about 0.17 SD. Further, the subgroup of young children with older siblings who, in turn, had between four and six years of education (when the CCT's education conditions become more binding due to high drop out in that range) had improvements of about 0.28SD, a very large effect. Effects are also larger for 0-3-year-old children, who are those that are starting to develop their language skills, as well as for households with older female siblings. More importantly, the effects seem to be mainly driven by the interactions of older female siblings with younger boys, which are those that one would expect in a model of females being more involved in the caregiving chores and in which boys present largest deficits in development. We interpret these results as the programme creating education spillovers between siblings. The results are comparable to direct ECD programme effects in Latin America, suggesting that a comprehensive policy which incorporates these spillovers can be a viable option for poverty reduction in the region.

Chapter 5 concludes the thesis by noting the contributions of the work, how the results together can form more innovative and effective policy, and finally, explores future avenues of research in the field opened by the research.

2 Identifying the Role of Nutrition in Developing Cognitive Skill in Young Peruvian Children

2.1 Introduction

Research that documents how early childhood conditions are vital in the development of human capital skills has gained traction in the economic literature over the last 20 years (see Almond and Currie [2010] for a thorough literature review). Poverty poses an important risk to children's development as its consequences limit a household's ability to provide nurturing environments, and respond to investment sapping shocks. This is especially pertinent in the early childhood, where children's brains are at their most malleable (Cunha and Heckman 2007). Barker (1995) introduces the biological mechanism behind this as the foetal origins hypothesis. Any external stressors to children in this early period, such as exogenous climate (Dercon and Porter 2014) and health shocks (Almond 2006) or resource deprivations as a result of poverty or cultural decisions (Schote, van den, Lindeboom 2012; Almond, Mazumder and Van Ewijk 2015; Majid 2015) can lead to poor health, unemployment, spells of incarceration, and an increased likelihood of risky behaviours. These outcomes have economic importance in terms of increases in the private and public costs of living.

A specific risk to children in the early stages of life is malnourishment. Malnourished children are deficient in key nutrients which affect brain development and diminishes their ability to interact with others and complete tasks. Households may then delay school enrolment for these children, disrupting the development of key relationships with peers. Skill gaps by nutritional status then begin to form into the middle and late childhood and perpetuate once children are in school as they lack the brain architecture and capacity to reach their full potential (Grantham-McGregor et al. 2007). Policymakers who aim to close these cognitive gaps for these children must know when investments are most effective. Most of the previous studies which show positive outcomes on skills development conclude that the early childhood is the most productive period, without explicitly comparing effects across the first five years of life (Barham, Macours, and Maluccio 2013; Grantham-McGregor et al. 2007; Walker et al. 2011). This question, 'when is investment in nutrition most productive in the first five years?', drives the present research.

In order to answer this question, a framework must be in place to explain and classify the evolution of the impacts of nutrition on cognitive development. Cunha and Heckman (2007) and Cunha, Heckman and Schennach (2010) introduce the two key terms that create this framework which drives this research; critical and sensitive periods of development. A critical period of development is when investment in a certain factor has a return at only one point in time. Sensitive periods are where investment in one period has a greater return on investment than other periods, but the return is not limited to that one period alone. There has been little focus on actually identifying these periods of development in longitudinal and experimental data when examining cognitive development (Van den Berg, Lundborg, Nystedt and Rooth (2014) is an example relating to health outcomes). Using data from the first three rounds of the Young Lives survey in Peru, this paper aims to estimate the existence of these periods in the first 5 years of life in terms of skills development at age 7 and 8. A production function for skills, based upon Todd and Wolpin (2003, 2007), is estimated in two ways; firstly, through simple models which provide lower bound estimates of the total, direct, and indirect effects of early nutrition on cognitive achievement, and second, through causal estimates which aim to provide more conclusive evidence of the existence of these periods.

The past literature in the field has come to mixed conclusions on the role nutrition has on developing cognitive skill and how the timing of nutritional investment can affect children's outcomes. In terms of direct effects, some papers find that nutritional disturbances in the early childhood do have significant impacts on children's skill development (Almond et al. 2014), while others show the opposite (Behrman and Duc 2014). In terms of timing, Glewwe and King (2001) state that the key investment period is age 12-24 months, while Barham et al. (2013) and others (Lopez-Boo and Canon 2014) state that investment leads to returns before age 1 and after age 4. The variety of results suggests that the research in the past has not successfully identified the true impact of nutrition on children's cognitive development, especially when using non-experimental methods. This paper is important because it aims to provide a better measurement of the effect of nutrition on cognitive development in non-experimental settings in two ways. First, it analyses associations between nutrition and cognitive ability in two important periods of childhood by obtaining the total effect and decomposing it into direct and indirect effects. Then, it builds on the instrumental variables methods in the field with the introduction of two new methods to increase the robustness of the estimates. This is the main contribution of the paper, as it highlights some of the challenges which previous papers have faced when identifying endogenous regressors, and offers a solution to the problem.

This research finds positive lower bound estimates of the effect of nutrition in both periods on cognitive ability. In terms of total effects, early childhood and preschool aged nutrition have an effect of approximately 0.1 SD on age 7 to 8 PPVT. These effects can be decomposed into direct and indirect effects through a life-course model (Tu, Tilling, Sterne, and Gilthorpe 2013), showing that the total effect of early childhood nutrition is half a direct effect and half an indirect effect. The presence of an indirect effect suggests that good nutrition in the early life can help unlock a child's economic productivity in future periods by building the cognitive architecture which complements later life development. In line with Popli, Gladwell, and Tsuchiya (2013), it can be concluded that there is a sensitive period of development in the early childhood as the effect size in the early childhood is larger than the preschool ages. The caveat to this is the fact that the two periods of nutrition are not significantly different than each other, with the main difference being that the estimate on early childhood nutrition is slightly more precise. Analysis of sub samples broken into stunted and non-stunted children yields larger effects for the stunted group, and no effect in the non-stunted subsample.

Next, the paper attempts to find causal estimates by using the traditional instruments in the literature, mother's height and birthweight, to identify the two periods of nutrition. This specification leads to imprecise estimates, motivating the use of a new methods to check the validity of these results, and possibly obtain strong enough identification to provide robust inference. The first method used is Lewbel's (2012) instrumental variable method which generates instruments from assumptions about heteroskedasticity in the data (see Emran and Hou 2013; Millimet and Roy 2015; Sabia 2007). Heteroskedasticity is likely to naturally occur in data as different values of one observation can lead to more variability in the outcome of interest. The method has not been used in the context of early childhood development to the author's knowledge. In this sense, it is an important methodological addition because it provides a source of exogenous instruments when there are no suitable sources in the data and can serve as a tool for researchers unable to run a wide-scale experiment. Additionally, the generated instruments can be used to test the validity of existing instruments through over identification tests. When using this method with the data in this paper, the traditional instruments are seen as valid (in an over identification sense) and there are no statistically significant impacts of nutrition investment on cognitive ability.

Another way to combat poor identification is to use weak instrument robust inference confidence sets. Weak identification leads to finite sample bias and poor inference. The confidence sets, based upon Anderson and Rubin's (1951) test of regressors, illustrate the range of possible parameter values in relation to each of the endogenous regressors and as a result is robust to this bias. Using both traditional and generated instruments, the weak identification robust methods show that nutrition in both periods simply is not identified. This is important because previously trusted instruments do not provide adequate identification of nutrition when identified in the same equation, casting a shadow on previous results in the field that do not take into account the strength of identification (beyond the Staiger and Stock [1997] rule of thumb). For example, the instruments used in Lopez-Boo (2009), Berhman and Duc (2014), and Glewwe and King (2001) do not work in this case, calling into question their wider applicability. In terms of this specific paper, it is impossible to make a conclusion on the causal impacts of nutrition on cognitive skill in this sample, much less determine critical or sensitive periods. This lack of conclusive evidence supports the idea that comprehensive investments which take into account improvements in nutrition along with parental and home help are more effective than solely focusing on nutrition, or any specific component.

The study makes use of the Young Lives data from Peru, which is particularly suited to answer questions on the development of children in a low to middle income setting. The data contains measurements of children's nutrition at age 6 to 18 months and 4-5 years old, as well as measures of cognitive skill in the mid-childhood. Peru is an interesting case study as it is a country that is undergoing a rapid economic renewal and serves as a bellwether for the rest of Latin America. Peru, and the region as a whole, has experienced decreases in poverty and economic inequality as wages and employment prospects have improved. The findings of the survey can serve as a reference point for Peru's place within the region as well as the region's place amongst the global economy (Georgiadis et al. 2017). Furthermore, the findings can help policymakers and researchers understand the conditions that are unique to those in poverty, and help extend solutions into other parts of the world.

The paper continues with a brief review of the past literature. It then continues by introducing the production function for skills, its implications, and the empirical specification that must be taken. The next step is to discuss the threats to identification, the instruments that have been used in the past, and to introduce the Lewbel IV to this context. Finally, the data is summarised and the results are presented and discussed.

2.2 Literature Review

The previous research on the effects of malnourishment in the early childhood has tended to focus on outcomes in the mid to late childhood, coming to mixed conclusions. Many papers show that achievement on cognitive assessments is predicted by the nutritional status of children at specific points of time across a range of different settings such as India, Peru, and Vietnam (Lopez-Boo 2009; Duc 2011; Helmers and Patnam 2011).

Malnourishment experienced in-utero and in the early stages of life has been shown to lead to worse health, labour market and skills outcomes for children as they age, no matter if it is the consequence of drought, poverty, religious reasons, or conflict. In terms of conflict, there is a modest literature focusing on the consequences of the Second World War on the long-term outcomes of children who were exposed to the war. Kesternich, Siflinger, Smith and Winter (2014) show that individuals who lived in areas of Europe during the war had lower self-reported health and lower education attainment, with the main channel of this effect being those who were malnourished in the early life. Scholte et al. (2015) and Jürges (2011) show that in-utero malnutrition as a result of the Dutch Famine Winter and Germany's post-war famine lead to higher hospitalisation rates, lower educational attainment, and worse occupational status. Lee (2014) provides similar results stemming from the Korean war in the 1950's. Outside of wartime conflict, Dercon and Porter (2014) show that famine exposure in Ethiopia led to children being shorter and accruing less earnings in the future (also see Woldehanna, Behrman and Araya 2013 for Ethiopian evidence). Almond et al. (2014) and Majid (2015) both illustrate the impacts of maternal fasting during Ramadan. Here, children who were in-utero during the fast period have lower achievement scores compared to other students. Behrman and Duc (2014) notes that the effects of nutrition investment may not extend beyond this in-utero stage, providing estimates which show that nutrition in the early childhood does not have a causal effect on cognitive ability for Vietnamese children.

The timing of nutrition measurement is important as well. Glewwe and King (2001), Crookston et al. (2011), Crookston et al. (2013), and Casale and Desmond (2016) all examine how investments at different points in the first five years of life predict better cognitive outcomes later in life, either through using absolute measurements or growth intervals. Glewwe and King (2001) provide estimates of causal impacts of nutrition as a result of children's growth between 12 and 24 months of age. Crookston et al. (2011) and Crookston et al. (2013) highlight the positive impacts in Peru and the other Young Lives countries of catch up growth[3] in the mid childhood (age 7-

[3] accelerated growth after lagging behind

8), but not the causal impacts. Casale and Desmond (2016) offer evidence, showing that South African children who recover from stunting before the age of 5 do no better on cognitive tests than those children who have not recovered from stunting. The authors use the first two years of life as their period of early childhood nutrition, rather than 6-18 months in the Young Lives data. The contrasting evidence from using different measurements in time for nutrition suggests that there may be an important characteristic of development that is missed in the early nutrition measurement in the Young Lives data.

There is also evidence which shows the importance of interventions which can lead to improvements in cognitive achievement. Hoddinott, Alderman, Behrman, and Haddad (2013) provides a review of the literature which shows the importance of interventions in the first 1000 days to reduce the incidence of stunting in children. Interventions play an important role in remediating these gaps by improving the accessibility of food and health services for those in rural and extremely impoverished settings. Barham et al. (2013) is an example, studying the effect of a conditional cash transfer programme in Nicaragua which led to huge achievement gains at age 10 for those who were treated before the age of 1 compared to those treated between 1 and 2, where there was little to no effect. Campbell et al. (2014) show that Carolina Abecedarian Project children who received home visiting before the age of five had larger health effects than those treated between the ages of 6 to 8. These results contrast the evidence in Glewwe and King (2001) and Casale and Desmond (2016) which point to growth between 1 and 2 as the key point of development in the first 1000 days of life. Lopez-Boo and Canon (2014) show positive impacts of nutrition in the preschool ages on cognition at age 8 in India for children who were enrolled in meals programme, supporting a conclusion of sensitive, rather than critical, periods in the first five years of life.

The literature review makes clear that the jury is out concerning when investment is most productive in the first 5 years of life and whether there are critical or sensitive periods of development. Part of this results from diverse data sets, while another part stems from the methodologies used by the authors. The mixed evidence beyond experimental and quasi-experimental studies suggests that there may be some error in inference when obtaining causal estimates, motivating the use of new methodologies (like the one presented in this paper).

2.3 Estimating the Production Function for Skills

In order to examine the impact of nutrition on skills, we must determine how skills are produced in the first place.[4] The production function for skills framework is attractive here because its flexibility allows for inputs over a period of time to have an impact on current skills, rather than just focusing on current stocks of investment. By explicitly separating the effects, the production function allows for research into understanding the role that each input plays in producing an outcome conditional on the other factors. Equation 2.1 shows the production function for cognitive skill for children at time t (which in this case is age 7-8 years old):

$$\theta_{it} = f(H_{i1000,}\mathbf{H}_{i,(t-l),}X_{it,}I_{i,(t-1),}, \mu) \tag{2.1}$$

θ_{it} is the revealed cognitive ability of the child, H_{i1000} denotes nutrition for the first 1000 days of life (important in the past literature), and $\mathbf{H}_{i,(t-l)}$ is a vector of lagged nutrition in other periods of time. The other inputs which are included in the production function are the household and individual characteristics, $X_{i,(t-l)}$, and $I_{i,(t-l)}$ which consists of different parental investments into the child, and are detailed further in section 3. The endowment u_i captures the genetic and heritable characteristics that a child is born with. In this study, *t-1* defines the preschool ages and *t-2* defines the early childhood. Contemporaneous effects are not used because of the possibility of reverse feedback between a child's nutrition and their cognitive ability and a timing assumption on when investment inputs kick in (Lopez-Boo 2009; Lopez-Boo and Canon 2014).

[4]See Todd and Wolpin (2003, 2007) and Cunha and Heckman (2007) for a more thorough discussion on the theoretical background.

The separation of nutrition into the first 1000 days of life and the other measurements beyond this is an evolution of the previous research which has emphasised the importance of the first 1000 days on the development potential of the child. Including $\mathbf{H}_{i,(t-l)}$ as a vector allows for more data (as available) on nutrition to be used if desired to determine how nutrition affects the cognitive development of a child over the life course, at the risk of more difficult inference if causal estimates are desired. In addition to this, the separation of nutrition allows for study into whether different periods of life are critical or sensitive in terms of the effectiveness of investment during these times. A critical period is identified as follows:

$$\frac{\partial \theta_{it}}{\partial H_{i,(t-l)*}} \neq 0; \ \frac{\partial \theta_{it}}{\partial H_{i,1000}} = 0 \tag{2.2}$$

A critical period of nutrition means that changing nutritional investment in one time period yields a non-zero effect on cognitive skill, while in all of the other periods it has no effect. In this illustrative example, investment must be directed to children older than 1000 days (3 years) if it is to be successful in boosting a child's cognitive ability.

For a sensitive period, the following holds:

$$|\frac{\partial \theta_{it}}{\partial H_{i,(t-l)*}}| > |\frac{\partial \theta_{it}}{\partial H_{i,1000}}| \neq 0 \tag{2.3}$$

Here, changing nutritional investment leads to greater absolute returns in one period compared to the others. In this example, improving nutrition in a later period of childhood has a larger effect on cognition than in the first 1000 days of life.

Determining if these periods exist is only possible if the production function has non-recursive nutrition inputs. Non-recursive inputs mean that health in one period is not 'produced' by nutrition in an earlier period (Popli et al. 2013). This assumption holds in the example where a child could be at a good level of nutrition in *t-1,* and become much worse off in period t as a result of a shock in the interim period that diminishes nutritional investment. When there is a recursive structure to the production function, health is wrapped up in multiple inputs which makes it impossible to disentangle the effects in each unique period.[5] Popli et al. (2013) also notes that identification of critical and sensitive periods is dependent on the frequency of data being used and the length of time it is meant to measure. If the frequency is low, then the measures are meant to account for a longer period of time and the likelihood of identifying a critical period is higher. When that critical period is separated by more frequent data, then it is more likely to be seen as multiple sensitive periods rather than as a critical period.

2.3.1 Empirical Specification

The original version of Equation 2.1 (without the inclusion of separate periods of nutrition) has been estimated as a linear function or as a structural model (Cunha, Heckman and Schennach 2010; Sanchez 2013). This paper follows the linear form of Equation 2.1, known as the cumulative specification because it uses an accumulation of measurements over time, rather than those in a single time period.

$$\theta_{it} = \beta_0 + \beta_1 H_{i,(t-l)} + \beta_2 H_{i1000} + \beta_3 I_{i,(t-1)} + \beta_4 X_{it} + \mu + \epsilon_{it} \tag{2.4}$$

Consistent estimation is attained by assuming that unobserved input effects are age-independent. That is, it must be assumed that the unobserved input effects do not have impacts outside of the difference in current age and the age of the child when the input occurred (Todd and Wolpin 2003, 2007). Additionally, the omitted inputs associated with the lagged variables must be orthogonal to the included inputs. This assumption is strong because it assumes that parents treat all their children the same and do not make investment decisions based on environmental and observational factors. The addition of the lagged component makes this more difficult because it means that parents are not using past knowledge in making investment decisions

[5] An earlier version of this paper included a value added structure with past test score. As a result, the past test score included impacts of health in its production function, meaning it was impossible to properly identify health effects

in their children. The descriptive evidence lessens this concern by using controls to capture all of these unobserved effects. Instruments can also be used to reinstitute the zero conditional mean condition for the explanatory variables of interest and obtain causal estimates of these effects.

Critical and sensitive periods of development for nutritional investment can be determined by comparing the coefficients β_1 and β_2 in the linear equation. Using the same hypothetical example from above, $|\beta_1| > |\beta_2| \neq 0$ would denote a sensitive period of nutrition investment for the middle period of nutrition, whereas $\beta_1 \neq 0;\ \beta_2 = 0$ would denote a critical period of nutrition investment in that same period.

Tu et al. (2013) discusses different empirical specifications which will affect the interpretation of the estimates in Equation 2.4. To be able to completely understand critical and sensitive periods, a conditional body size model is used to identify total effects. The conditional body size model includes the initial measurement of nutrition and a conditional body size measure, which is the residual of a regression of future measurements of nutrition on the previous measurements. The residual can be viewed as the difference between the actual nutrition of the child and the predicted nutrition in the early childhood (Tu et al. 2013, p. 1332). For this paper, the residual comes from a regression of height for age in the preschool ages on the same measure in the early childhood, and is expressed in Equation 2.5.

$$H_{it-1} = \alpha + H_{it-2} + \epsilon_{it} \tag{2.5}$$

The conditional body size measure is by construction uncorrelated with nutrition in the early childhood, lessening endogeneity concerns and fitting better with the assumption of non-recursive nutrition. In comparison, using the previous measures of height for age, as is, is the life-course plot model. In this model, the coefficients are interpreted as the direct impacts of each period of nutrition on cognitive ability rather than the total effect of that period. This model is weak when discussing critical and sensitive periods of development because it ignores the indirect effects that different periods of nutrition have on cognition. For example, the direct associations between preschool nutrition and cognitive ability may be larger than those between early childhood nutrition and cognitive ability, but the total association of early childhood nutrition are larger because of stronger indirect effects. It is necessary to include this information to ensure that the results are being properly understood and explained.

2.3.2 Addressing Endogeneity

Identification of the variables of interest, early childhood and preschool aged nutrition, is subject to a number of factors which can bias the estimates. Firstly, it is difficult to establish the direction of causation between nutrition and cognition because of unobserved factors. A child could be smart because they are healthy and have better brain function, or because they have a better knowledge of what needs to be done to stay healthy. Since children are under the age of 5 in this study, it is likely that the parents are "choosing" nutritional investments for the child rather than the child "choosing" for themselves. A greater risk of endogeneity comes from unobserved investment changes which stem from parent's observations of the health and cognitive ability of their child. This will bias the estimates upward or downward depending on why parents change the level of investment. Reinforcing investments, or those that are a response to better revealed outcomes for a child, lead to upward bias, as the investment would seem to be the driver when in reality it was the child's existing level of skill. Estimates of nutrition would be larger when in reality the improved cognition was as result of increased investment due to the child performing well. "Compensating" investment means that parents try to equalise levels of cognitive ability, shrinking observed gaps rather than rewarding the child for being smart (Lopez-Boo 2009). The previous literature (Glewwe and King 2001; Lopez-Boo 2009; Duc 2011; Lopez-Boo and Canon 2014) has provided evidence of a downward bias, as the magnitude of the estimates increase once instrumental variables estimators are used.

2.3.3 Identification concerns

This study faces two main identification concerns. The first is weak or under identification of the endogenous regressors due to the rank condition, $E(z_i x_i')$, being weakly satisfied (when this expectation is close to zero), or not satisfied at all (when it is zero). Under identification leads to inconsistent estimates whereas weak identification leads to finite-sample bias in the IV estimator. The bias is in the direction of the OLS estimates because the instrument does not have much predictive power through the endogenous regressor. The weaker the correlation between the instruments and the endogenous regressors, the higher the sample bias (Hahn and Hausman 2003). Anderson and Rubin (1951) and Kleibergen (2005) both provide methods which account for weak identification. These methods drop the rank condition and focus on the weak exogeneity condition and hypothesis tests of the coefficients of interest. Point estimates are discarded and confidence sets (in the case of two endogenous regressors) are created under an assumption of linearity. Each method is tested for different maintained and alternative hypotheses. The Anderson-Rubin test uses the weak exogeneity condition in the null and alternative hypotheses, while Kleibergen's K and the CLR statistics test whether $\beta = \bar{\beta}$. In these two cases, the maintained hypothesis is that weak exogeneity holds.

The next identification concern is the fact that the two measures are taken from different periods in time. In both age periods, the different unobserved effects of the child, family, and environment can uniquely effect both health and cognitive ability. The unobserved component of the early childhood can also be related to the measurement of preschool aged nutrition, bringing in a time series component. In order to have a full rank variance-covariance matrix, the instruments must be relevant for each period of time. For example, instruments that are only measured in Round 2 of the survey could serve as valid instruments if they are uncorrelated with the error in the next period. However, the correlations are considered spurious, because a future feature, be it a policy or different measurement of health, should not have an impact on data from the past.[6]

[6] Unless of course it was expected, and therefore it would be endogenous!

2.3.4 Instrumental Variables

Valid instruments must meet the exclusion and relevance assumptions for both of the endogenous variables. The selected instruments must be external to the model in the sense that they affect the level of skills through height-for-age and not on their own, satisfying the zero conditional mean assumption, $E(Z\varepsilon) = 0$. Additionally, they must not be correlated with the unobserved factors. Finally, the instruments must be highly correlated with the endogenous regressors so that changes in the instrument can work through the endogenous regressor, rather than unobserved factors. It can be the case that data in longitudinal studies do not meet these requirements as variation may arise endogenously through unobserved household decisions. In addition, there is not always scope to study the impacts of an idiosyncratic shock or a policy shift to create this exogenous variation.

Two instruments for health and nutrition that have been used in past studies are birthweight and mother's height. Birthweight has been used (Duc 2011; Lopez-Boo 2009; Hoddinott and Kinsey 2001) as it captures health directly after the in-utero period. It should not be correlated to cognitive ability because it captures an absolute measure of health and nutrition rather than height for age, which is inherently a comparative measure. Concerns still exist over its endogeneity as socioeconomic status may play a role in its determination. One way of combatting this is to use a residual component of birthweight, motivated by Behrman and Duc (2014). The residual captures the impact of shocks on the child in the gestation period. Mother's height is another candidate as an instrument as it could indicate a mother's opportunity cost of working as well as early life disease susceptibility (Glewwe and Jacoby 1995). Mothers who are taller may have higher returns from working, which in turn could affect their child's cognitive ability. Disease susceptibility could be genetically linked amongst family members, therefore identifying this specific component of a child's health (Glewwe and Jacoby 1995, p. 161). The two instruments should not have an effect on the cognitive ability of the child, because they are unrelated to the decision making of the parents after birth. Genetically speaking then, the mother can transmit these health traits that she has developed to the child, but the environmental changes (very few) have no effect on outcomes at age five. In addition to these variables, food prices for oral rehydration solutions, rice, potatoes and milk in the communities are used in line with Glewwe and King (2001), as well as food price shocks. An additional set of instruments is necessary to try and improve identification.

2.3.5 Lewbel Instrumental Variables

Lewbel (2012) introduces a method which is useful when there is lack of exogenous variation or concern with instrument validity. The method generates exogenous instruments from assumptions about heteroskedasticity in the data. Heteroskedasticity is a common feature in data as the conditional variance can change across different individuals and households. The generated instruments are useful in just identified models to improve the efficiency of the IV estimator and to test questionable instruments through over identification tests. To motivate the method, a generalised form of the estimating equation is found in Equation 2.6. Equations 2.7 and 2.8 show the reduced form equations of the two endogenous regressors in a triangular Lewbel setup.

$$\theta_{i,t} = \beta_1 H_{i,(t-1)} + \beta_2 H_{i,1000} + \beta_3 X_{i,t} + u_{i,t} \tag{2.6}$$

$$H_{i,t-1} = \beta_2 X_{i,t} + \varepsilon_{2i,t} \tag{2.7}$$

$$H_{i,1000} = \beta_3 X_{i,t} + \varepsilon_{3i.t} \tag{2.8}$$

For the framework to create suitable instruments, there are three identifying assumptions. Firstly, there must be heteroskedasticity in the first-stage equations to generate the instruments (it is not necessary in the second stage). Higher moment conditions are used as a result, reducing the efficiency of the estimator. Assuming that heteroskedasticity is present is abnormal only in the sense that it becomes a formal requirement, rather than something that needs to be addressed through adjusting standard errors.

$$cov(X_{i,t}, \epsilon_{ji}^2) \neq 0;\ j = 2, 3 \tag{2.9}$$

Lewbel (2012) suggests using the Breusch-Pagan (1979) heteroskedasticity test to identify if heteroskedasticity exists and determine its scale. The Breusch-Pagan test searches for heteroskedasticity across the entire first stage equation, rather than the individual components, which can make it difficult to identify whether the 'right' heteroskedasticity is present in the data. The 'right' heteroskedasticity in this case is heteroskedasticity which comes from the error term, rather than from the unobserved factor that affects both the endogenous regressors and the dependent variable. The test makes a strong assumption of normal errors. The heteroskedasticity in the reduced form equations is exploited to create new variables which can be used as instruments. The construction of the weighting matrices for identifying the parameters of interest is flexible enough to allow for both two stage least squares and GMM estimation.

The next assumption is that it is necessary for the instrument to be weakly exogenous to the reduced form and structural errors.

$$cov(X_{i,t}, \epsilon_{ji}) = 0;\ j = 2, 3 \tag{2.10}$$

$$cov(X_{i,t}, u_i) = 0 \tag{2.11}$$

Lastly, cross-equation homoskedasticity is required and is one of the fundamental assumptions of the method.

$$cov(X_{i,t}, u_i \epsilon_{ji}) = 0;\ j = 2, 3 \tag{2.12}$$

Cross-equation homoskedasticity means that there is homoskedasticity in the system as a whole, just not necessarily in the individual equations on their own. This condition ensures point estimation because the covariance matrix must be a diagonal matrix and is essential for the Breusch-Pagan test to work.[7] Set identification is possible if the condition does not hold. For more details, see Equation 15 in Lewbel (2012) and the associated discussion.

Now that the identifying assumptions have been met, the instruments can be generated.

$$\ddot{Z}_i = (X_{i,t} - \bar{X}_{i,t})\hat{\epsilon_{ji}};\, j = 2, 3 \tag{2.13}$$

In the above, the X variables are demeaned and multiplied by the residuals of the reduced form equations. X can comprise of all of the exogenous variables, or a selection of a few (perhaps just the ones with heteroskedasticity with respect to the endogenous variables). Because ε_2 and ε_3 are not observable, a sample moment is taken from the reduced form equations. If heteroskedasticity is high, the generated instrument will be stronger and $cov(\ddot{Z}_i, \hat{\varepsilon}_{ji})$ will not be zero or close to zero. The generated variables are mean zero which ensures that they are exogenous to the other included regressors. They are then used as instruments either on their own or in conjunction with other existing instruments. The inclusion of more instruments improves efficiency at the cost of less precision. More variance in the regressors signifies that the instruments are better able to identify the endogenous regressor because the correlation between the instruments and the endogenous regressors is proportional to the covariance between these instruments and the squared errors. Why? Heteroskedasticity in the error process means that as the value of the dependent variable changes, the residuals change in value as well. The scale of this change is used in the creation of instruments Z, so elements are shared and a proportional relationship is present.

[7] See STATA help file on ivhettest

Figure 2.1: Young Lives Peru Interview Sites

2.4 Description of Young Lives Peru

The data used in this paper comes from the Young Lives survey that is administered by the University of Oxford and supported by the United Kingdom Department for International Development (DFID) among others. The Young Lives dataset provides a comprehensive look at the conditions faced by children in poor households in lower to lower-middle income countries. The study began in 2001 and interviews take place on three year intervals for a total of 15 years in Ethiopia, India,[8] Peru and Vietnam. The selection and interview process follows standard ethical norms, which insures the anonymity of the sample and is transparent. In the Peruvian study, the interviews take place in 20 randomly chosen communities called "sentinel sites" scattered across Peru (seen in Figure 2.1). In Peru, the survey is nationally representative of all but the top 5% of the income distribution (Escobal and Flores 2008).

[8]in Andhra Pradesh and Telangana

An older and younger child cohort were selected to take part in the study so that it would be possible to take into account the different stages of life. This study focuses on the younger cohort which consisted of 2000 children that were born in 2001 and 2002, giving a comprehensive view of a child's development into young adulthood. The younger cohort includes measurements from the key stages in the developmental cycle of a child including ages 6-18 months old, 4-5, and 7-8 years old. Attrition from Round 1 to Round 3 is low, with 3.75% of the sample not being interviewed in each of the rounds.[9] The surveys collect a wide array of data from the interviewed households and children. Children whose first language is Quechua, the primary indigenous language of Peru, are not included because the Spanish translated cognitive test is not an adequate measure of their cognitive ability.[10] Children who had past nutrition measurements greater than the absolute value of five are dropped, as these measurements are infeasible (Georgiadis et al. [2017] supports this). Additionally, children who have not completed the PPVT test are removed leaving a sample of 1471 children. Of these, 160 children do not have measured birthweight in the sample. Removing these children completely from the study would increase the concerns of selection bias, so their values are imputed.[11] The imputation procedure comes from Behrman and Duc (2014) and regresses birthweight on mother's weight, wealth index from the first round of the survey, the number of weeks a child was born premature, and dummy variables which denote a child's health in comparison to other children in the community, self-reported by the mother. These children are poorer than the other children on average and were more likely to be born at home rather than in a hospital. A t-test shows that measured birthweight and the imputed values are not significantly different from each other.

2.4.1 Variable Description

The production function for skills in Equation 2.1 has five key components; cognitive ability, two nutrition measures, indicators for investment and household controls.

[9] See Outes-Leon and Dercon (2008) for more details.
[10] This operation has been carried out in most of the studies to date
[11] The appendix shows the results of the main specifications less these children.

The outcome measure, cognitive skill, is a flexible term, and the true measure of these skills is as opaque as their definition given in multiple different avenues of the research. There is not an assessment or question that can truly measure cognitive skill. In the Young Lives data, the best available proxy is the Peabody Pictoral Vocabulary Test, which is an exam of receptive vocabulary (Cueto and Leon 2012; Dunn, Padilla, Lugo and Dunn 1986; and Dunn and Dunn 1997). The PPVT test has been used as a measure of cognitive skills in many papers using the Young Lives country studies (see Lopez-Boo 2009; Crookston et al. 2013) and other studies (Paxson and Schady 2010). Two other cognitive tests were taken for the Young Lives survey; the Early Grade Reading Assessment (EGRA) and a maths test. These tests focus on understanding and problem solving, and can provide a fuller picture of how nutrition affects development. All three measures are standardised by age.

The endogenous explanatory variables of interest are the lagged nutrition variables measured by height-for-age z-score. Height-for-age is an indicator for health and nutritional investment in a child up until the time of measurement. It is measured in all three rounds of the survey at the international standard set by the World Health Organization (2009). According to the WHO, the height-for-age z-score is a measurement for malnutrition as it measures the length of a child at their current age on a scale of measurements for children around the world. This makes it possible to compare nutrition across different countries. The ages when nutrition is measured in the survey are important as well. As Almond and Currie (2010) and the results of Barham et al. (2013) have shown, the first 1000 days of life are key for improving child outcomes later in life. While the Young Lives data does not specifically contain a measurement of nutrition at 1000 days, the early childhood nutrition measurement is capturing the halfway point, and can be seen as an imperfect proxy for these 1000 days. The mid-childhood period at age 4-5 captures many key milestones such as weaning and school entry.

The next part of the production function is the investment indicator. This paper combines ideas in Lopez-Boo (2009) and Helmers and Patnam (2011) and uses a number of different household and individual characteristics, such as a child's frequency of seeing their father, whether the father is alive, if the child plays more than two hours a day, and if the child is responsible for doing tasks or chores around the house. The investment variable is the sum of all of these variables once they have been standardised mean 0 and variance 1. This ensures that higher values of the index mean higher levels of investment into the child while also accounting for the multi-dimensional nature of parental investment.

The controls take into account individual and household characteristics, aiming to account for possible channels nutritional investment could travel through to affect a child's cognitive ability. In terms of individual level controls, age in months, gender and birth order are used. For the caregivers and households, the size of the household, the education level of the primary caregiver, the height of the mother, and the wealth of the household are included. The wealth status of a household is measured on a 0 to 1 scale, which combines household quality, consumer durables, electricity, and quality of sanitation, fuel and drinking water. The wealth index and caregiver's education are both highly correlated with nutrition in the pre-school ages, and are transformed to limit the effects of multicollinearity. In the former, the logarithm of the wealth index is constructed. In the latter, a dummy variable is constructed with values of 1 denoting that the caregiver has more than 8 years of education. Other geographical measures such as living in an urban or rural area are used as well as whether a child is in a public or private school. Lastly, regional fixed effects are included based on whether the household is located in the coastal, jungle or mountainous regions of Peru.

2.5 Results

The analysis begins with descriptive statistics and a discussion of the makeup of the samples being used. Table 2.1 shows the descriptive statistics for the working sample and for the sample including the children whose mother tongue is Quechua. These children are able to recall on average 62 words on the PPVT test. Their measured nutrition in general decreases from round 1 to 2, leading to an increase in stunting percentages at this time. This is likely due to children being weaned, putting them at risk for malnourishment in low income households. The percentage of the sample reported as being stunted at least once in their lives is 35%. The children at time t are about 8 years old, and the sample is balanced in terms of gender. Households have five members on average, are predominately urban, and report wealth on the upper half of the Young Lives wealth index (better housing, water quality, etc.). Caregivers (mainly mothers) have on average 8.7 years of formal education. Focusing on the parental investment variables, the majority of children see their fathers on a daily or weekly basis, their fathers are living, and they spend adequate time playing. The largest difference between the the samples is in the raw PPVT score. This is to be expected because the test is administered in Quechua.

Table 2.1: Summary Statistics

	mean	sd	mean	sd
PPVT 7-8 years old	62.19	15.91	61.63	16.29
Height for Age 4-5 years old	-1.36	1.06	-1.39	1.11
Height for Age 1-2 years old	-1.09	1.20	-1.10	1.29
Child height (cm) 4-5 years old	105.25	5.99	104.22	8.29
Child height (cm) 6-18 months years old	71.83	4.50	71.79	4.62
See father at least weekly	0.79	0.41	0.79	0.41
Is child's father still alive?	0.99	0.09	0.99	0.11
Does household chores	0.67	0.47	0.67	0.47
Plays at least two hours per day	0.97	0.17	0.97	0.17
Does task for household	0.11	0.31	0.11	0.31
Mother's Height R1	150.30	5.43	150.28	5.44
Female	0.49	0.50	0.49	0.50
Age of child (mths)	95.31	3.60	95.33	3.63
Household size	5.26	1.86	5.27	1.86
Caregiver's education level	8.71	4.19	8.61	4.22
Urban	0.81	0.39	0.80	0.40
Wealth Index R3	0.58	0.20	0.58	0.20
Has been stunted at least once in their lives	0.35	0.48	0.35	0.48
Private School	0.22	0.41	0.21	0.41
Observations	1471		1552	

Note: Columns 1 and 2 are the working sample.

Columns 3 and 4 include children dropped because their mother tongue is Quechua

2.5.1 Total, Direct, and Indirect Associations

The research in the past (Almond and Currie 2010) has discussed the first five years of life as a critical period of development. Additional empirical literature (Barham et al. [2013] and others) extends this conclusion by emphasising the importance of the first 1000 days of life. An advantage of the Young Lives data is that it has measured children at two different points in the first 5 years of life. This allows for a test of when nutritional estimates have the largest effect on cognitive ability in the mid childhood. While the two periods are not perfectly timed to test the first 1000 days hypothesis, the results are still informative on the productivity investments at different points in time. Building on the previous literature, Georgiadis et al. (2017) show that direct and total effects (which should differ) can be obtained by using two different specifications; the life-course plot model for the former and the conditional body size model for the latter. The results presented here illustrate the evolving role of nutrition in the first 5 years of life on children's early cognitive outcomes.

Table 2.2 shows the estimates of the linear production function for skills (Equation 2.2) in terms of the conditional body size model (Columns 1 and 2), which uses the residual from a regression of later measures of nutrition on the previous measures, and the life-course plot model (Columns 3 and 4), which uses height for age when children are 6-18 months old and 4-5 years old. In terms of total effects, both early childhood and preschool nutrition have a 0.1SD association with a child's cognitive ability at age 8. According to Popli et al. (2013), the early childhood is a sensitive period because the coefficient is slightly larger than its preschool aged counterpart. It is a weak conclusion though, because it is a marginal and statistically insignificant difference. Analysis of the controls introduces a few key points of note. Firstly, the wealth index and caregiver's education are significant predictors of a child's ability to no surprise. Richer and smarter households are better able to provide the necessary tools to develop cognitive skill in their children as well as not having to face the stressors which are inherent in impoverished households. Interestingly, the home inputs score is insignificant.

The life-course plot model is then used to decompose this total effect into direct effects. The association between preschool nutrition and cognitive ability is unchanged (as expected) in this model, signifying that this period only has a direct effect on

Table 2.2: Associations of Nutrition with Cognitive Skill at Age 7-8

	Conditional Body Size		Life-Course Plot	
Conditional Body Size, age 4-5	0.321***	0.102***		
	[0.2,0.4]	[0.05,0.1]		
Height for Age 4-5 years old			0.321***	0.102***
			[0.2,0.4]	[0.05,0.2]
Height for Age 6-18 months old	0.219***	0.102***	0.0513	0.0494*
	[0.2,0.3]	[0.06,0.1]	[-0.01,0.1]	[-0.0006,0.10]
Parental Investment in Round 3		0.0626		0.0626
		[-0.03,0.2]		[-0.04,0.2]
Mother's Height R1		-0.00685		-0.00685
		[-0.02,0.001]		[-0.02,0.002]
Female		-0.108**		-0.108*
		[-0.2,-0.009]		[-0.2,0.0005]
Age of child (mths)		0.00505		0.00505
		[-0.005,0.02]		[-0.006,0.02]
Household Size		-0.000337		-0.000337
		[-0.02,0.02]		[-0.03,0.03]
Caregiver has completed up to the secondary education		0.255***		0.255***
		[0.2,0.3]		[0.2,0.3]
Urban		0.205*		0.205*
		[-0.004,0.4]		[-0.02,0.4]
Mother's age at birth		0.0140***		0.0140***
		[0.005,0.02]		[0.004,0.02]
Log Wealth of the Household		0.717***		0.717***
		[0.5,0.9]		[0.5,0.9]
Birth Order		-0.0973***		-0.0973***
		[-0.1,-0.05]		[-0.2,-0.04]
Private School		0.238***		0.238***
		[0.1,0.4]		[0.09,0.4]
Observations	1471	1471	1471	1471
R-squared	0.146	0.370	0.146	0.370

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.
Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.
Column 1 contains no controls, Column 2 adds household and individual controls and regional fixed effects
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

cognitive ability at age 8. Early childhood nutrition has a much smaller effect on cognitive ability, with the direct association now less than half of the total association at 0.049 SD. This effect is still not statistically different than the effect of preschool aged nutrition however. The fact that the direct effect is larger in this case is not a surprise, since it comes from a measurement which is closer in time to the measurement of cognitive ability. An interesting finding is that the proportion of early childhood nutrition which travels through indirect means to impact children's cognitive ability is nearly half of the total effect of early childhood nutrition on cognitive ability. This suggests that improvements in early childhood nutrition could build the framework on which future skills are developed. Previous literature which has only focused on direct associations has missed out on identifying possibly over half of the total effect of early childhood nutrition on cognition. These results add more credence to the importance of early childhood investment on cognitive ability.

2.5.2 Causal Estimates

It is expected that the above estimates of the impacts of nutrition in the early childhood and preschool ages, even with the controlling variables, suffer from endogeneity. The sign of the bias is dependent on how investments into the child are made. The previous literature (Lopez-Boo 2009) states that it should be a downward bias because parents base their investment decisions on their judgement of their child's nutrition, and provide compensatory investment to boost their child's chances. Table 2.3 shows the initial estimations for this specification following the same format for the two models as Table 2.2 used. In both models, early childhood nutrition is positive and significant while preschool nutrition is insignificant and negative. Both coefficients increase in absolute terms as expected. The conditional body weight model is more precisely estimated than the life course plot model, but overall it is still imprecise. In comparison, the direct effects from the life course plot model are larger than the total effects in the conditional body size model as a result of the imprecision in the data. The change in sign on the preschool aged effect provides additional evidence of the effects of imprecise estimates. The effect is insignificantly different than zero but the large standard errors and some bias leads to a large decrease in the estimated coefficient. A negative impact of nutritional investment has little economic significance and goes against the previous literature which emphasises the importance of nutrition in the early ages.

There are two factors which could be causing the inefficiency of the estimators in the IV specification. On one hand, high correlations between the variables of interest and these two variables could lead to multicollinearity and unreliable estimates. On the other hand, the selected instruments may not be properly identifying both endogenous variables. Identifying two endogenous regressors is difficult, and is made further difficult by the time component. Any weakness in identification may not be apparent from the weak and under identification test statistics. One way to solve these problems is to use weak identification robust inference and construct confidence sets rather than obtaining point estimates. Weak identification robust inference is based upon loosening the assumption that the endogenous regressors are strongly identified.

Table 2.3: Causal Estimates of Nutrition on Cognitive Skill at Age 7-8

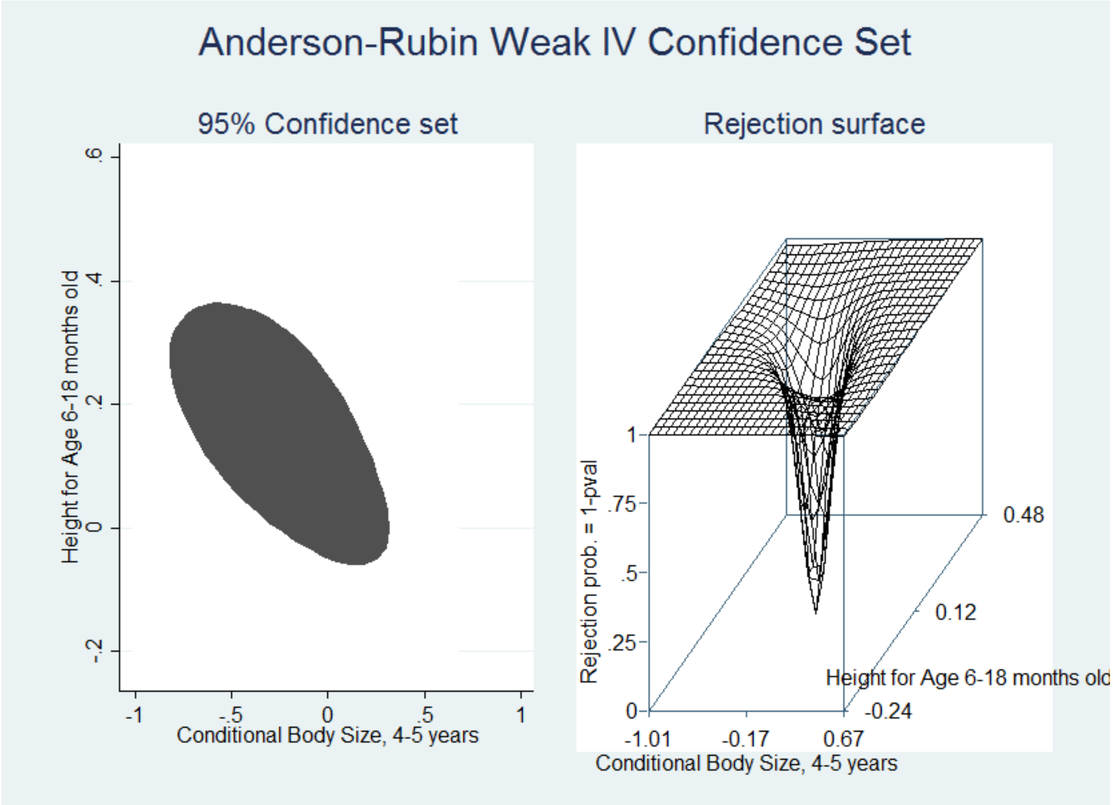
	Conditional Body Size		Life-Course Plot	
Conditional Body Size, age 4-5	0.180	-0.169		
	[-0.10,0.5]	[-0.5,0.1]		
Height for Age 4-5 years old			0.180	-0.169
			[-0.10,0.5]	[-0.5,0.1]
Height for Age 6-18 months old	0.259***	0.117*	0.165	0.205*
	[0.1,0.4]	[-0.003,0.2]	[-0.1,0.4]	[-0.03,0.4]
Parental Investment in Round 3		0.0706		0.0706
		[-0.02,0.2]		[-0.02,0.2]
Female		-0.150**		-0.150**
		[-0.3,-0.03]		[-0.3,-0.03]
Age of child (mths)		0.0177		0.0177
		[-0.005,0.04]		[-0.005,0.04]
Household Size		-0.00567		-0.00567
		[-0.04,0.02]		[-0.04,0.02]
Caregiver has completed up to the secondary education		0.296***		0.296***
		[0.2,0.4]		[0.2,0.4]
Urban		0.217**		0.217**
		[0.005,0.4]		[0.005,0.4]
Mother's age at birth		0.0195***		0.0195***
		[0.01,0.03]		[0.01,0.03]
Log Wealth of the Household		0.785***		0.785***
		[0.6,1.0]		[0.6,1.0]
Birth Order		-0.116***		-0.116***
		[-0.2,-0.06]		[-0.2,-0.06]
Private School		0.265***		0.265***
		[0.1,0.4]		[0.1,0.4]
Observations	1471	1471	1471	1471
R-squared	0.129	0.326	0.129	0.326
Under-ID	8.838	9.744	8.838	9.744
Weak-ID	10.82	12.74	10.82	12.74
J Test	0	0	0	0

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.
Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.
Instruments are mother's height and birthweight of child
Column 1 contains no controls, Column 2 adds household and individual controls and regional fixed effects
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Inference based on confidence intervals allows for the lack of precision to be baked into any conclusions made from the estimation process. The presence of two endogenous regressors makes inference more difficult as the selected instruments may identify one regressor better than another.

In the standard IV framework, weak identification can be identified with the Kleibergen-Paap (2006) test statistic. According to Staiger and Stock (1994), test statistics of 10 or higher are sufficient for unbiased point estimates. While the initial causal results in Table 2.3 meet this criteria (a test statistic of 12.7), the impreciseness and possibility of bias means that robust estimation methods should be used. The confidence sets, which are shown in Figure 2.2, are set at a significance level of 0.05.

Figure 2.2: Confidence Set with Birthweight and Mother's Height as Instruments



The confidence set highlights the weakness in identification, as it is impossible to rule out a null effect. Using projection based inference (a conservative form of inference), it is possible to find bounds for the confidence set. In the case of early childhood nutrition, the projection based confidence interval is (-0.14, 0.33), while preschool aged nutrition is (-0.80, 0.25). Negative values can likely be ruled out, because any investment into a child should at the very least have no effect. These bounds in concert with visual inspection seem to point towards there being larger effects (ranging from just above 0 to approximately 0.2) in the early childhood compared to the preschool ages. However, no strong conclusion can be made based solely on the results contained in Table 2.3 and Figure 2.2.

Table 2.4 adds the additional instruments, such as food prices, adverse events and the residual of birthweight are used. Glewwe and King (2001) offer an example of identifying multiple endogenous regressors using food prices and rainfall data as their instruments. The authors justify the use of food prices as instruments by explaining how variation is created by relative income and cost for each household. These include food prices (Column 1) whether the household experienced an exogenous shock of any

Table 2.4: Additional IV Regressions

	Food Prices	Shocks	Birthweight Residual
Conditional Body Size, age 4-5	-0.0865	-0.139	-0.100
	[-0.3,0.2]	[-0.4,0.1]	[-0.4,0.2]
Height for Age 6-18 months old	0.0714	0.121**	0.0822
	[-0.06,0.2]	[0.0001,0.2]	[-0.03,0.2]
Parental Investment in Round 3	0.0616	0.0699	0.0672
	[-0.04,0.2]	[-0.02,0.2]	[-0.03,0.2]
Female	-0.0797*	-0.147**	-0.137**
	[-0.2,0.01]	[-0.3,-0.02]	[-0.3,-0.02]
Age of child (mths)	0.00940	0.0168	0.0118
	[-0.01,0.03]	[-0.007,0.04]	[-0.009,0.03]
Household Size	-0.00567	-0.00514	-0.00465
	[-0.03,0.02]	[-0.03,0.02]	[-0.03,0.02]
Caregiver has completed up to the secondary education	0.274***	0.289***	0.289***
	[0.1,0.4]	[0.2,0.4]	[0.2,0.4]
Mother's age at birth	0.0222***	0.0188***	0.0189***
	[0.01,0.03]	[0.010,0.03]	[0.010,0.03]
Urban	0.309***	0.215**	0.215**
	[0.08,0.5]	[0.004,0.4]	[0.002,0.4]
Log Wealth of the Household	0.764***	0.775***	0.778***
	[0.5,1.1]	[0.6,1.0]	[0.6,1.0]
Birth Order	-0.121***	-0.113***	-0.115***
	[-0.2,-0.07]	[-0.2,-0.06]	[-0.2,-0.06]
Private School	0.305***	0.261***	0.259***
	[0.1,0.5]	[0.1,0.4]	[0.1,0.4]
Observations	942	1471	1468
R-squared	0.347	0.335	0.347
Under-ID	11.04	10.40	10.30
Weak-ID	31.18	7.677	15.02
J Test	.	4.914	0

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.
Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.
Initial instruments are mother's height and birthweight
Column 1 adds food prices (price of oral rehydration solutions, potatoes, rice, noodles and milk)
Column 2 adds if the household experienced an exogenous shock
Column 3 uses the residual birthweight component
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

kind (Column 2). Column 3 then combines mother's height with a residual which measures the cumulative shocks to a child when they were in the womb, based upon the method used in Behrman and Duc (2014). The residual component of birthweight aims to reduce endogeneity concerns in birthweight by measuring the accumulation of exogenous shocks while in-utero. These results do little to improve the inference, and motivate the necessity of a new method to try and improve the validity of the results.

2.5.3 Generated Instruments

The Lewbel (2012) procedure is useful in situations where it is difficult to obtain reliable instruments in the data based on theory. Breusch-Pagan tests of the first stage regressions with the selected sets of instruments are used to determine if heteroskedasticity is present. Rejecting the null hypothesis of the test means that heteroskedasticity is present and the instruments should be able to identify the endogenous regressors. Table 2.5 shows the results of the preferred estimation (Column 1) and two additional specifications using different sets of variables to be used for creating instruments. In the specification with the entire set of controls this leads to more instruments than clusters. The controls are partialled out in the second stage regression by way of Frisch-Waugh-Lovell (Frisch and Waugh 1933; Lovell 1963) to square the variance-covariance matrix. As was the case before, each specification is weakly identified, thus confidence sets are used. In each of the regressions, Hansen's over identifying restrictions test indicates that the traditional instruments are valid. Figure 2.3 shows the 95% K and AR confidence sets and rejection regions for each set of generated instruments.

The first column of Table 2.5 shows the estimation results where only mother's height is used as the variable to create the instrument. In this specification, the Breusch-Pagan test statistics are 2.7 for early childhood nutrition (which is at the 10% significance level) and 25.93 for preschool aged nutrition. The key takeaway from the results is that the endogenous regressors are only weakly identified. As a result, the point estimates are similar to the OLS for preschool aged nutrition and larger for early childhood nutrition. Only early childhood nutrition is statistically significant. The first figure in Figure 2.3 emphasises the weak identification as it is impossible to rule out most results. The confidence set includes the OLS estimates from Table 2.2, as well as larger values which could include unbiased point estimates based on the prior knowledge. A lower bound can be placed on early period nutrition at approximately -0.22 SD. In terms of preschool nutrition, the projection-based confidence interval contains all realistic values and as a result, accurate inference is impossible. From the rejection surface, it is apparent that the effects are robust for many confidence levels, confirming that the width of the set stems from a lack of identification, not of power. Using the estimates contained in this paper as well as the past results, it seems that the early childhood nutrition will have a stronger effect than preschool nutrition, but this is a difficult conclusion to make.

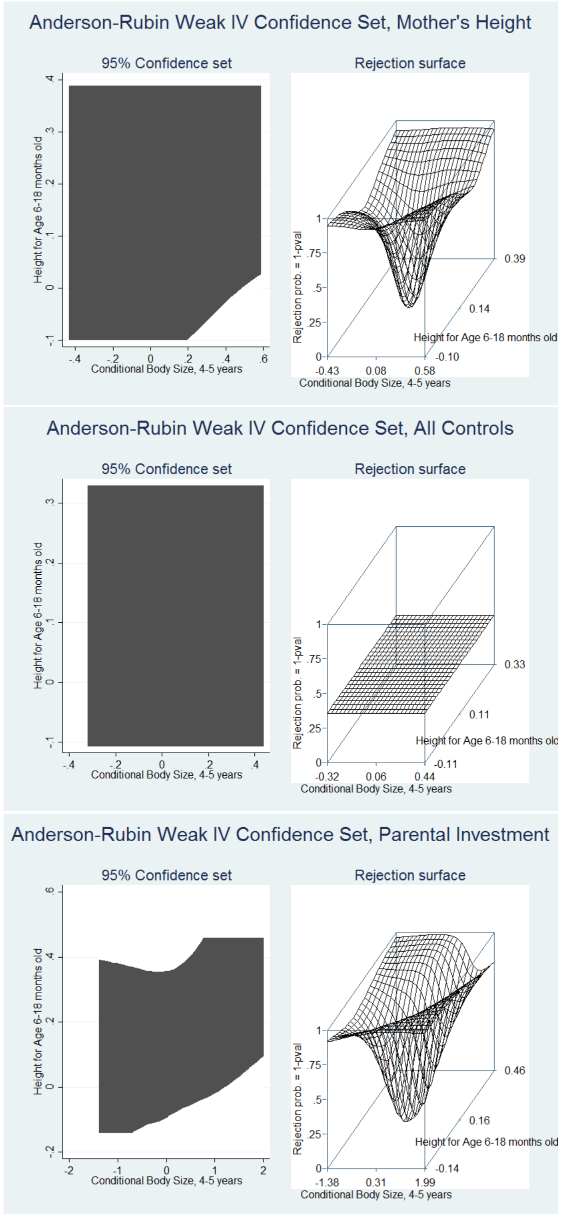
Table 2.5: Regressions of Nutrition on Cognitive Skill at age 7-8, using generated instruments

	Mother's Height	All controls	Parental Investment
Conditional Body Size, age 4-5	0.0758	0.0589	0.309
	[-0.2,0.3]	[-0.1,0.2]	[-0.5,1.2]
Height for Age 6-18 months old	0.145**	0.111**	0.159**
	[0.02,0.3]	[0.002,0.2]	[0.010,0.3]
Parental Investment in Round 3	0.0660		0.0609
	[-0.03,0.2]		[-0.03,0.2]
Mother's Height R1	-0.00828		-0.0155
	[-0.02,0.004]		[-0.05,0.01]
Female	-0.118**		-0.0852
	[-0.2,-0.02]		[-0.2,0.03]
Age of child (mths)	0.00983		0.00123
	[-0.01,0.03]		[-0.03,0.03]
Household Size	-0.000731		0.00393
	[-0.03,0.03]		[-0.02,0.03]
Caregiver has completed up to the secondary education	0.255***		0.217***
	[0.2,0.3]		[0.07,0.4]
Urban	0.206*		0.195*
	[-0.002,0.4]		[-0.007,0.4]
Mother's age at birth	0.0138***		0.00854
	[0.004,0.02]		[-0.01,0.03]
Log Wealth of the Household	0.710***		0.643***
	[0.5,0.9]		[0.4,0.9]
Birth Order	-0.0953***		-0.0771*
	[-0.1,-0.04]		[-0.2,0.008]
Private School	0.238***		0.213**
	[0.1,0.4]		[0.04,0.4]
Observations	1471	1471	1471
R-squared	0.367	0.0218	0.344
Under-ID	4.918	.	7.603
Weak-ID	2.595	.	3.994
J Test	0.215	.	0.00341

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.
Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.
Column 2 partials out the controls using Frisch-Waugh-Lovell
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The next step is to show how changing the instruments changes the identification of the two nutrition effects. The second figure uses all of the controls and investment measures as the variables to create instruments from. While heteroskedasticity is present for both reduced form equations (Breusch-Pagan statistics of 18.4 for the early period, and 57.9 for the preschool period), there is still an under identification problem with the endogenous regressors. This creates an unintelligible confidence set and rejection surface where almost every value nutrition in the early and middle period is possible. The last figure shows the case where only the parental investment index is used and there is no heteroskedasticity. Again, the endogenous variables are under identified and the resulting confidence set behaves poorly. Overall, these results indicate that even by using advanced methods, proper identification of nutrition is not attained in this data.

Figure 2.3: Confidence Sets with generated instruments



It is worth briefly discussing the heteroskedasticity that is present in this data. While the Breusch-Pagan test statistics indicate that there is heteroskedasticity in the data, it could be the case that there needs to be a very high level of heteroskedasticity for strong identification to occur. Remember, the strength of the generated instrument is determined by the scale of heteroskedasticity in the data. The poor identification in these results suggest that the scale of heteroskedasticity here is not of the right magnitude. While disappointing in the scope of this paper, it can serve as an example for the literature on using the methodology when the conditions are not perfect.

2.5.4 Robustness Checks

In addition to the above analysis, there are a number of robustness checks that support the results from above. Considering the possibility that the PPVT score does not capture all of the components of cognitive skill, additional language and math tests, which are taken by the child during the surveying process, are used as the outcome variable. Table 2.6 shows these results, with OLS regressions in the first two columns, and IV regressions in the second two. For brevity, only the conditional body size model is used. In both cases, the effect sizes are slightly smaller than the PPVT results, but the results follow the same trend. In the IV regressions, the coefficients behave in the same way as they do in the preferred results, although they are now insignificant.

Table 2.6: Results using Math and Language Tests as Measures of Cognitive Skill

	OLS		***IV***	
	EGRA	Math	EGRA	Math
Conditional Body Size, age 4-5	0.0828***	0.0714***	-0.0354	-0.0382
	[0.03,0.1]	[0.03,0.1]	[-0.5,0.4]	[-0.4,0.3]
Height for Age 6-18 months old	0.0769***	0.0800***	0.102	0.0504
	[0.05,0.1]	[0.06,0.1]	[-0.05,0.3]	[-0.07,0.2]
Parental Investment in Round 3	0.0742	0.0783*	0.0788	0.0788*
	[-0.03,0.2]	[-0.01,0.2]	[-0.02,0.2]	[-0.005,0.2]
Mother's Height R1	-0.00199	-0.00454		
	[-0.01,0.01]	[-0.01,0.005]		
Female	0.0236	-0.145***	0.00291	-0.158***
	[-0.04,0.09]	[-0.2,-0.05]	[-0.08,0.08]	[-0.3,-0.05]
Age of child (mths)	0.00185	0.00321	0.00905	0.00529
	[-0.009,0.01]	[-0.01,0.02]	[-0.02,0.04]	[-0.02,0.03]
Household Size	-0.0137	0.00103	-0.0160	-0.00124
	[-0.04,0.009]	[-0.03,0.03]	[-0.04,0.010]	[-0.03,0.03]
Caregiver has completed up to the secondary education	0.140***	0.239***	0.156**	0.259***
	[0.04,0.2]	[0.1,0.3]	[0.01,0.3]	[0.1,0.4]
Urban	0.307***	0.140	0.312***	0.145
	[0.1,0.5]	[-0.1,0.4]	[0.1,0.5]	[-0.1,0.4]
Mother's age at birth	0.0196***	0.0109**	0.0217***	0.0138**
	[0.01,0.03]	[0.0007,0.02]	[0.009,0.03]	[0.001,0.03]
Log Wealth of the Household	0.418***	0.492***	0.442***	0.531***
	[0.3,0.6]	[0.3,0.7]	[0.2,0.7]	[0.3,0.7]
Birth Order	-0.103***	-0.0851***	-0.110***	-0.0957***
	[-0.1,-0.07]	[-0.1,-0.04]	[-0.2,-0.07]	[-0.2,-0.04]
Private School	0.152***	0.349***	0.162***	0.362***
	[0.04,0.3]	[0.2,0.5]	[0.05,0.3]	[0.2,0.5]
Observations	1469	1468	1469	1468
R-squared	0.199	0.246	0.190	0.238
Under-ID			9.739	9.392
Weak-ID			12.60	11.48
J Test	0	0	0	0

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.

Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.

Columns 3 and 4 use Mother's Height and Child's Birthweight as Instruments

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.7 shows the sample cut into male and female, and urban and rural. These specifications have similar results to what is seen in Table 2.2 save for the case of female children where the effects are indistinguishable.

Table 2.7: Results for Gender and Urban/Rural Subsamples

	Male	Female	Urban	Rural
Conditional Body Size, age 4-5	0.113***	0.0793*	0.0999***	0.114**
	[0.04,0.2]	[-0.003,0.2]	[0.04,0.2]	[0.02,0.2]
Height for Age 6-18 months old	0.0743***	0.144***	0.0905***	0.141**
	[0.02,0.1]	[0.06,0.2]	[0.04,0.1]	[0.02,0.3]
Parental Investment in Round 3	-0.0239	0.134*	0.0281	0.161
	[-0.1,0.10]	[-0.004,0.3]	[-0.04,0.10]	[-0.1,0.4]
Mother's Height R1	-0.00244	-0.0127**	-0.00148	-0.0270**
	[-0.01,0.010]	[-0.02,-0.002]	[-0.01,0.007]	[-0.05,-0.0003]
Female			-0.0869	-0.176
			[-0.2,0.02]	[-0.5,0.10]
Age of child (mths)	0.00122	0.00924	0.00225	0.0110
	[-0.01,0.02]	[-0.005,0.02]	[-0.007,0.01]	[-0.01,0.04]
Household Size	-0.0153	0.0131	0.00684	-0.0191
	[-0.05,0.02]	[-0.02,0.05]	[-0.02,0.03]	[-0.09,0.06]
Caregiver has completed up to the secondary education	0.215***	0.295***	0.186***	0.536***
	[0.09,0.3]	[0.1,0.4]	[0.08,0.3]	[0.4,0.7]
Urban	0.227*	0.200		
	[-0.02,0.5]	[-0.04,0.4]		
Mother's age at birth	0.0155***	0.0127**	0.0152***	0.0181***
	[0.004,0.03]	[0.0009,0.02]	[0.004,0.03]	[0.006,0.03]
Log Wealth of the Household	0.659***	0.763***	0.655***	0.714***
	[0.4,0.9]	[0.5,1.0]	[0.4,0.9]	[0.4,1.0]
Birth Order	-0.105**	-0.0896***	-0.102***	-0.104**
	[-0.2,-0.02]	[-0.1,-0.04]	[-0.2,-0.05]	[-0.2,-0.01]
Private School	0.204**	0.275***	0.223***	0.727
	[0.02,0.4]	[0.2,0.4]	[0.09,0.4]	[-0.3,1.7]
Observations	748	723	1193	278
R-squared	0.324	0.432	0.255	0.361

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.
Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 2.8 looks specifically at the effects of stunting. As the descriptives showed, 35% of the sample was measured as stunted at least once in the first 5 years of life. Column 1 provides evidence of a sensitive period of development in the preschool ages with significantly larger effect sizes than in the full sample. It is unclear whether these estimates are total or direct effects, making it difficult to come to any conclusion on sensitive or critical periods. The result is interesting because it gives evidence of heterogeneity across different levels of nutrition and highlights how limiting extreme undernutrition may lead to improvements in cognition. Columns 2 and 3 cut the sample into groups of children who were reported as being stunted once or not. The results in Column 2 are particularly noteworthy because only preschool aged nutrition has an effect on cognitive ability, and it is small. It could be that this is evidence of diminishing returns of cognitive skill production from nutrition investment once children are no longer stunted.

Table 2.8: Results focused on Stunting

	Stunting	Not Stunted	Stunted
Conditional Body Size, age 4-5		0.0538*	0.202***
		[-0.003,0.1]	[0.1,0.3]
Height for Age 6-18 months old		0.0207	0.231***
		[-0.05,0.09]	[0.2,0.3]
Stunted in R2	-0.168***		
	[-0.2,-0.10]		
Stunted in R1	-0.156***		
	[-0.3,-0.05]		
Parental Investment in Round 3	0.0593	0.0252	0.0822
	[-0.04,0.2]	[-0.09,0.1]	[-0.1,0.3]
Mother's Height R1	-0.00307	-0.00217	-0.0149**
	[-0.01,0.004]	[-0.01,0.007]	[-0.03,-0.002]
Female	-0.117**	-0.109*	-0.112
	[-0.2,-0.02]	[-0.2,0.01]	[-0.3,0.05]
Age of child (mths)	0.00375	-0.00315	0.0179*
	[-0.008,0.02]	[-0.02,0.009]	[-0.001,0.04]
Household Size	-0.000338	0.00554	-0.00714
	[-0.03,0.02]	[-0.02,0.04]	[-0.04,0.03]
Caregiver has completed up to the secondary education	0.265***	0.275***	0.202***
	[0.2,0.4]	[0.2,0.4]	[0.1,0.3]
Urban	0.207*	0.167	0.212**
	[-0.01,0.4]	[-0.07,0.4]	[0.0002,0.4]
Mother's age at birth	0.0165***	0.0154***	0.0102
	[0.008,0.03]	[0.005,0.03]	[-0.007,0.03]
Log Wealth of the Household	0.721***	0.745***	0.688***
	[0.5,0.9]	[0.5,1.0]	[0.5,0.9]
Birth Order	-0.106***	-0.0728**	-0.112***
	[-0.2,-0.06]	[-0.1,-0.010]	[-0.2,-0.06]
Private School	0.246***	0.234***	0.248**
	[0.1,0.4]	[0.1,0.4]	[0.01,0.5]
Observations	1471	963	508
R-squared	0.365	0.278	0.364

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.
Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.
Column 2 contains children who have not been stunted in the first 5 years of life while Column 3 contains those who were.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The last robustness checks take into account concerns over measurement of children's nutrition. Height for age z score is the generally accepted measure of health and nutrition in the early childhood, but there is still discussion on whether it is the best measure of nutrition in children. To take into account this concern, Table 2.9 runs the primary regressions using child's absolute height rather than height for age. The results lean more conclusively towards the early childhood being a more important period for the development of cognitive skill. In the OLS regressions, there is a larger gap between the associations between nutrition in the two periods and cognitive ability. When the IV estimates are considered, the preschool aged estimates are closer to zero while the early childhood estimates are significant and positive. Overall though, the results are largely consistent across both measures.

Table 2.9: Results using Child's Height as Measure of Nutritional Status

	OLS		IV	
Conditional Body Size, age 4-5	0.0743***	0.0220***	0.0301	-0.0385
	[0.06,0.09]	[0.01,0.03]	[-0.04,0.1]	[-0.1,0.02]
Child's Height at 6-18 months old (cm)	0.0440***	0.0358***	0.105**	0.0553*
	[0.03,0.06]	[0.02,0.05]	[0.01,0.2]	[-0.002,0.1]
Parental Investment in Round 3		0.0646		0.0591
		[-0.03,0.2]		[-0.04,0.2]
Mother's Height R1		-0.00664		
		[-0.01,0.001]		
Female		-0.0495		-0.0246
		[-0.1,0.05]		[-0.2,0.1]
Age of child (mths)		-0.0255**		-0.0613*
		[-0.05,-0.005]		[-0.1,0.006]
Household Size		-0.000441		-0.00562
		[-0.02,0.02]		[-0.03,0.02]
Caregiver has completed up to the secondary education		0.251***		0.304***
		[0.2,0.3]		[0.2,0.4]
Urban		0.196*		0.223**
		[-0.01,0.4]		[0.003,0.4]
Mother's age at birth		0.0136***		0.0191***
		[0.004,0.02]		[0.009,0.03]
Log Wealth of the Household		0.706***		0.807***
		[0.5,0.9]		[0.6,1.0]
Birth Order		-0.0957***		-0.116***
		[-0.1,-0.05]		[-0.2,-0.06]
Private School		0.231***		0.285***
		[0.1,0.4]		[0.1,0.4]
Observations	1471	1471	1471	1471
R-squared	0.170	0.371	0.0461	0.315
Under-ID			5.775	9.112
Weak-ID			4.953	11.27
J Test	0	0	0	0

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.
Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.
Each specification follows the same trend: Column 1 has no controls;
Column 2 adds household and individual controls as well as regional fixed effects
Mother's Height and Child's Birthweight are used as instruments in Columns 3 and 4
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Tables A.1 and A.2 present results with different samples and different measures of the conditional body size model. The results are largely the same across the different specifications.

2.6 Discussion and conclusions

This paper aimed to develop a better idea of how socioeconomic status travels through nutrition to effect cognitive ability in the mid-childhood and identify when nutritional investment is most productive in the development of skills. In this study, nutrition is measured in the early childhood (6 to 18 months) and preschool ages (4-5 years). The previous literature has either focused either solely on associations between nutrition in these two periods of time (Crookston et al. 2013; Georgiadis et al. 2017) or on identifying the impacts of early childhood nutrition on cognitive ability through the use of multiple instruments (Behrman and Duc 2014). To the author's knowledge, Glewwe and King (2001) is the only other paper which looks to identify multiple periods of nutrition with causal methods. In addition to this, the paper aims to identify if there are critical or sensitive periods of development, showing where nutritional investment is most productive for mid-childhood skills development.

The overall body of evidence in the paper points to the early childhood being a sensitive period of development at the very least. OLS estimates of the associations between nutritional investment and cognitive achievement find effect sizes of a 10 percent increase in PPVT score after a 1SD improvement in nutrition in both periods of development. The total effect can be broken down into direct effects by using a life course plot model, showing that the total effect of early childhood nutrition is half a direct effect and half an indirect effect. It is important that these indirect effects are identified because they highlight channels of development that focus more on increasing the productivity of children rather than directly improving a later achievement metric. Without them, there is a sort of omitted variable problem. Being healthier and well nourished means that you can focus more, are more able to deal with stress and adversity, and to tap into your resource base to perform better on assessments. Being unhealthy early in life could mean that your body is developing at a slower rate than those who are healthy. Slow development means that when the child enters school, he or she lacks the ability to make the most of the teaching and interactions with peers that they are exposed to on a daily basis. The results are similar to the previous literature (Crookston et al. 2013; Georgiadis et al. 2017). Crookston et al. (2013) shows a similar trend of the total effects between the two periods being indistinguishable (although the authors use a longer period of unpredicted growth; see Table A.2 for comparison). Georgiadis et al. (2017) has slightly different results, either because of the selected sample or because they use a path model to estimate their relationships rather than the linear model used here. The larger estimates for 'ruralites' compared to their urban counterparts highlights some of the disparities between localities. In poorer households, nutrition can make up for poorly stimulating environments. These associations highlight the intergenerational transfer of poverty for these children as the effects extend beyond the wealth status of the household. Improvements in nutrition may close some of these gaps, but there still are larger factors in play.

The IV estimates build on these results and the past literature by attempting to obtain causal estimates of the impact of nutritional investment in these two periods on cognitive ability. Nutrition in this case is considered endogenous because households may make investment decisions based on their judgement of their child's health. Parents invest in their children by either substituting inputs (helping their children along) or compensating through additional investment. The previous literature (Lopez-Boo 2009) has stated that compensatory investment is more likely, thus the associations in Table 2.2 are taken to be lower bound effects. As parents see their children lagging behind in terms of nutrition, they invest more resources into the child to compensate for their poor nutrition, and the apparent effect of nutrition is reduced. Nutrition in these two periods is instrumented first by birthweight and mother's height. The full specification in each model suggests that there is in fact a critical period of development in the early childhood, but the imprecision in the pre-school aged nutrition coefficient suggests poor identification. While the weak identification statistic meets the Staiger and Stock (1994) rule of thumb, weak instrument robust methods present large confidence intervals that make inference difficult. Using the assumption that nutritional investment will never hurt cognitive achievement, the evidence from the confidence regions suggest that the early childhood has a larger causal relationship with cognitive achievement than the pre-school ages.

Since these estimates are imprecisely estimated, a few additional steps are taken to try and improve the results. Additional instruments, such as the residual component of birthweight (Behrman and Duc 2014) and food prices (Glewwe and King 2001), are used to improve identification to little success. Food prices poorly identify the endogenous regressors in this case, while the residual component of birthweight shows little difference to the absolute birthweight measure. Next, additional instruments are generated through heteroskedasticity in the data. The most useful generated instrument comes from the heteroskedasticity associated with mother's height. This may act as a proxy for a woman's opportunity cost of working. Children's mothers who work may have healthier children not withstanding their economic status and this isolates the effect of nutrition. Even as it is the 'best' instrument, it still performs poorly. The additional instruments add little in the way of efficiency or towards improved inference.

These results display two significant caveats to the existing literature which are present in the descriptive and causal analysis. Skills development is a complex and dynamic process which relies on many different inputs. Understanding how poverty affects development is important for policymakers, as exploiting the channels it flows through for policy can be a cost-effective way to close skill gaps. In the case of the descriptive literature, estimates of associations are important to highlight the trends in a sample when it comes to the impacts of nutrition on cognitive ability. Far too many papers treat these results as causal effects, or near causal effects, when in reality they are not. This paper contributes to the understanding of nutrition by defining the effects total, direct, and indirect. Highlighting total, direct and indirect effects shows the many facets of development that nutrition may affect. Direct effects are those in which improved nutrition is related to improvements on an achievement measure, while indirect effects are those which improve the productivity of later life investment. This separation is an important dividing point with the existing literature.

The paper's result of poor identification of the endogenous nutrition regressors highlights highlights pitfalls in estimating causal effects of nutrition on cognitive ability in longitudinal studies. To combat these issues, it provides methods such as weak identification robust estimatiors and the Lewbel (2012) generated instruments method, that can be used to improve inference when model performance is suboptimal. The selection of instruments is important when trying to identify these effects. When using instruments that have been used in the past, the estimates are imprecise. This is because of the difficulties in identifying nutrition in two different periods of time for the same child when using longitudinal data. The Lewbel (2012) method provides a way to overcome this difficulty, but only when there is a high level of the right heteroskedasticity. It must be used with caution as the instruments add little to the quality of the estimator without this condition. Future research can continue to use this method to see if it is viable to use for inference in the developmental setting. As was noted above, the results demonstrates the importance in using weak instrument robust methods of inference. In the two endogenous regressor case, the confidence set and rejection contours indicate the weakness of the model even as it meets the criteria that has been set in the past literature for valid inference. The confidence sets complement the point estimates by displaying how the the true parameter values interact with each other, while in the rejection contours the power of the model is displayed. The body of evidence presented here suggests that the early childhood is more important, but that the weak power of the model makes strong inference impossible.

The attempt to estimate causal impacts through the use of instrumental variables adds to the sparse literature on evidence of causal impacts of nutrition on later life outcomes and highlights the pitfalls of causal inference based on imprecise estimates, even if test statistics support the validity of the model. The point estimates using a number of different instruments generally point toward early childhood being a critical period of development. The use of robust inference illustrates that statistically significant and imprecise point estimates are not sufficient for causal interpretation. The weak instrument robust confidence sets show that a number of different, competing interpretations cannot be ruled out. As such, no confident conclusion can be made. Use of these methods is important in future research, because it ensures accurate inference that can push the literature base forward. Even as this paper ends without a firm conclusion on the impacts of early childhood nutrition, it serves as an important guide for future research on identifying the causal channels through which socioeconomic status travels.

3 The importance of family, friends and location on the development of human capital in mid-childhood and early adolescence

"Train a child in the way he should go, and when he is old he will not turn from it."–Proverbs 22:6.

3.1 Introduction

Children's development is a complex process driven by many factors. Poverty, through material deprivations, can cause stressors to the household which hamper development. Poor children then fall behind well-developed counterparts with these gaps persisting into adulthood, affecting an individual's ability to be successful. Studies in the past (Schady et al. 2015; Lopez-Boo 2016) have focused on these socioeconomic differences in the realm of cognitive skills, where these gaps stabilise after age 10. The literature highlights potential channels through which deprivation travels in low income households to cause these gaps, such as malnourishment, food insecurity and other shocks to diminish skills development (Crookston et al. 2011; Crookston et al. 2013).

While cognitive skills are important measures, non-cognitive skills or "psychosocial competencies" are also essential components of a child's later life success. These competencies, also known as soft skills or socioemotional skills, can comprise multiple different domains of "skill" or personality traits. They are a set of ideas, beliefs or traits which enable individuals to tap into their cognitive skills and build a framework for success. They include, but are not limited to, an individual's self-esteem, self-efficacy, locus of control, pride and agency. Almlund et al. (2011) link economics, psychology and personality together by defining personality as something formed as a response to different inputs. The authors provide economic models to how these traits can be developed, the stability of traits over time, and the predictive power of these traits on positive later life outcomes. This study, along with others, show that traits stabilise as children age beyond the early adolescence (Trzesniewski et al. 2003; Cobb-Clark and Schurer 2012), motivating interventions to boost stocks of these traits while they are flexible.

This paper focuses on self-esteem and pride, two similar, but separate constructs which help children complete tasks and be good citizens. Individuals with higher stocks of these competencies have been shown to have higher probabilities of labour market success (Cubel et al. 2014). In this paper, self-esteem is a general measure of one's own feelings about themselves, while pride is more directed at pinning down how one values themselves. It can be described in terms of one's feelings on their own worthiness and competency in skills (Mruk 2009) not on their ability to complete tasks. Almlund et al. (2011) show that self-esteem is linked to neuroticism, one of the Big 5 personality traits in psychology. Rosenberg (1965) and Rosenberg et al. (1971) describe and measures one's feelings about their worth and their ability in a general sense. According to Branden (1994), its wider importance stems from it enabling an individual to access their characteristics to be successful. Therefore, individuals with higher self-esteem can tap into their cognitive skills and translate them into success in their work.

In the Young Lives setting, pride is a subset of self-esteem based upon the Rosenberg (1965) scale. Dercon and Krishnan (2009) explain this measurement as something more focused on the environment that a child is raised in, and the child's view of their worth in that context, rather than as a generalised measure of self-esteem (Dercon and Krishnan 2009, p. 4). As such, it is less related to the negative extreme of pride, arrogance or egoism. Lea and Webley (1997) expand this idea, describing pride as an expression of self-esteem, and being related to something that is your own. From an economic psychology perspective, the authors explain that proud individuals perform better, as individuals want to give a good account of themselves on tasks. Previous Young Lives studies which have examined psychosocial competencies has used pride as a proxy of self-esteem because of its close relation to self-esteem when the generalised questionnaire is not used (Dercon and Krishnan 2009; Dercon and Sanchez 2011; Favara and Sanchez 2017). This literature also shows how poverty is a determinant of pride in the early life, and how it can affect behaviour in the teenage years. Obtaining estimates for both of these psychosocial competencies provides a test of the quality of pride as a proxy of self-esteem in the past. Additionally, using pride advances the existing literature forward in Young Lives by showing how pride changes as children age into the early adolescence.

It is important, therefore, to understand how an individual's self-esteem and pride in themselves develop over time, as higher levels of each can improve an individual's labour market success and lower public and private costs. This paper aims to describe and explain the role of socioeconomic status in children's psychosocial competencies in the later developmental stages of childhood, and then estimate how policy can close gaps. The period studied here, age 6 to 12, is unique because it is one of the first attempts to illustrate socioeconomic gaps in these competencies as children age beyond the early childhood. The first half of the age range captures these competencies just after children enter school, so it is likely that any gaps at this point are not being driven by school factors alone. The latter half captures entry into middle-school and the early beginnings of income generating activities. This is a key point in the psychological development of a child, as they are faced with new challenges relating to their maturation and their growing stature in their environment. The data in the study is taken from the younger cohort of the Young Lives study in Peru, providing a useful longitudinal panel to study across the early life course. The study administers a set of questions on these psychosocial competencies, relationships to peers and parents, and other household and individual level data to the index children and their younger siblings. In terms of self-esteem, this paper uses the Rosenberg (1965) Self-Esteem scale for the first time in the Young Lives sample. Exploiting the sibling component of the survey is an important development for the literature, as it allows for household fixed effects to be removed, improving identification. Examining whether equalising inputs could close gaps is an important task because it identifies whether policies which aim to improve relationships and community aspects are viable options to reduce socioeconomic inequality and urban/rural disparities. Haney and Durlak (1998) motivate this task, as they show that interventions aimed at improving self-esteem in the early adolescence can be successful as they take advantage of the stabilisation of these traits.

The paper comes to three important findings. First, non-parametric analysis shows the presence of stable socioeconomic gaps in children's self-esteem, pride, and their relationships with peers and parents. These gaps motivate the parametric analysis of wealth effects and wealth gradients. In terms of locality, there are small, but statistically insignificant gaps for each measure except the parent-child relationship, which is large. This difference highlights a disparity in the way children relate to their parents in settings where work may be more prevalent in the younger ages.

Second, it is shown that the most important determinant of children's self-esteem and pride is the quality of their relationships with their parents and peers. Altogether, relationships with peers and parents accounts for between 50 to 80 percent of the variation in self-esteem and pride depending on which outcome is considered. The closeness of the measures confirms that pride has been a useful proxy for self-esteem in the past. The individual effect of the parent-child relationship is large compared to the previous literature which measures the direct transfer of specific psychosocial competencies, suggesting that existing estimates are lower-bound estimates of the role parents play on children's development. Siblings difference estimates support this conclusion, showing estimates similar in magnitude once household unobservables are removed. Analysis across the distribution of children's stocks of self-esteem and pride show variation in the determinants of these competencies. Children with different levels of self-esteem and pride derive their current level form different sources, an interesting outcome.

Lastly, equalising the parent-child relationship for socioeconomic status and locality can lead to gaps between children closing significantly. In terms of socioeconomic status, the effect is largely centred in pride, where the relationship can close 21% of the gap in pride. For locality, improving the parent-child relationship can close nearly 80% of the explained gap, a huge amount. This result serves as an indicator of a key difference in the situation between urban and rural households, and highlights an important channel for policy to exploit to close gaps. In doing so, the policy must recognise this unique situation to ensure not to introduce new problems to the dynamic in rural households.

The paper continues as follows. Section 2 takes a more thorough look at the previous literature in the field linking personality to economics, the development of psychosocial competencies, and their impact on future outcomes. Section 3 describes the data, providing data analysis and the smoothed differences in psychosocial competencies. Section 4 focuses on the empirical specification while Section 5 considers the results. The paper concludes in Section 6 with a discussion of the contribution of the paper in the wider literature.

3.2 Literature Review

There are two strands of economic research focusing on pride, self-esteem, and 'non-cognitive' skill. One focuses on the determinants of these competencies and skills through models, empirical analysis of longitudinal data, and experimental settings centred on after-school programmes. The other focuses on the larger economic consequences of individuals with higher or lower levels of these competencies.

An individual's self-esteem and pride can be formed through a variety of different pathways. Benabou and Tirole (2003) and Wydick and Darolia (2011) take economic approaches at modelling the relationship in principal agent type problem. In each case, children develop self-esteem and pride based upon how their parents, the principal, treats them. For example, Wydick and Darolia (2011) use measures such as how parents react to poor performance in school through hiring tutors, or whether a child had a car purchased for them, as different types of behaviours which can affect self-esteem. An easier to apply theory in this setting is the production function for skills developed by Todd and Wolpin (2003, 2007) which treats skills as something developed through investment in different inputs (such as health, education). The production function can be estimated to determine the relative predictive (or causal) power of certain inputs across the life course.

Almlund et al. (2011) build on these discussions by introducing personality to the economics literature by defining personality traits as responses to different inputs which can be modelled. These personality traits, namely the Big 5 personality traits of openness, conscientiousness, extraversion, agreeableness, and neuroticism are related to constructs like self-esteem (through neuroticism), linking psychological constructs to constructs more familiar to economists. Kautz et al. (2014) suggest that these traits are best measured through actively doing some sort of activity, such as answering a questionnaire or committing a crime (Kautz et al. 2014, p.16). The Rosenberg (1965) Self-Esteem scale falls into this category, as does the Young Lives pride subset. The importance of these personality traits on later life outcomes is explained through a wealth of applied evidence.[12] Heckman et al. (2006) show that non-cognitive skills, notably neuroticism, improves an individual's ability to acquire new skills and their productivity. Heckman and Rubenstein (2001) show worse employment prospects for GED graduates. Daly et al. (2015) shows that longer spells of unemployment for those with low self-control. In the case of the Young Lives literature, Favara and Sanchez (2017) study how low levels of pride lead to risky behaviours in Peru.

The authors also discuss the possibility of modifying these traits over time, providing an array of intervention examples leading to positive later life gains in personality traits (Heckman et. al 2010). A related topic is stability. Research in neurological science has shown that personality traits and psychosocial competencies are flexible for longer periods than cognitive skills because of the development of the prefrontal cortex (Dahl 2004; Knudsen 2004). Trzesniewski et al. (2003) show that self-esteem is at its most flexible point in the childhood. It then increases in stability as children age into adulthood where it peaks and then decreases until the end of life. The early adolescent period is noteworthy, as it denotes a rapid change in a child's environment as they face new challenges with peers and their understanding of their own self. Cobb-Clark and Schurer (2012) support this finding using the Big Five personality traits, showing that traits are stable beyond the age of 25. Robins and Trzesniewski (2005) provide a caveat to this conclusion, showing that while the mean-level of self-

[12]The literature in this field is wide and varied. This is a very brief overlook of the literature related to the topics studied in this paper.

esteem itself is unstable during this period, an individual's rank order amongst others is stable over time.

Branching from this literature is a rich examination of the determinants of pride, self-esteem, and other traits. In terms of the Young Lives literature, Dercon and Sanchez (2011) show that stunting in children predicts lower amounts of pride and agency. Dercon and Krishnan (2009) show that socioeconomic status can affects children's psychosocial outcomes in multiple ways. Parents, filled with the stressors of living in a low-income household likely would have lower stocks of pride which then appear in their interactions and behaviour with their children. It would also dampen the investment into the child, as there is not enough income to buy development promoting goods, or be able to send their child to school or after-school programmes. Outside of the Young Lives data, Heckman et al. (2006) show that schooling can be an important source of higher levels of personality traits. Krishnan and Krutikova (2013) evaluate an after-school arts and sports programme which led to children having higher levels of reported self-esteem. Relationships with parents and peers and good home environments to develop in are important characteristics to consider. Early in life, the home is where we develop fundamental learning and social skills which complement skills development in the future. Later in life, home should be the protective unit from the outside stressors of life, be it schooling, jobs, or relationships with others. Socioeconomic status plays a role in shaping the characteristics of this environment, but it is not a necessary condition for a good household. Poor households can practice good parenting while rich households can have parents who are not present in their children's lives.

In terms of the direct role of parents, particularly mothers, Georgiadis and Hermida (2014) and Georgiadis (2017) both show that parents have direct effects on their children's psychosocial competency development. In the first case, the authors show that there is a significant positive relationship between a parent and child's psychosocial wellbeing at age 7 and 8. In the second case, the analysis is extended out to age 12, showing that these positive associations grow with age. Anger (2012), Loehlin (2011) and Duncan (2005) all show intergenerational correlation coefficients of around

0.11-0.24 standard deviations between parent's psychosocial competencies and their adolescent children's in different settings in Germany and the United States. The gap in the research is how the role of parents, peers and the home environment evolve over time, and how relationships in the early adolescence can stem inequalities which form early in life.

In terms of closing gaps, much of the existing literature has focused on reducing wealth gradients rather than closing gaps in the mean-level of pride or self-esteem. In the early childhood, Fernald et al. (2012) show that socioeconomic gaps in early childhood development indicators in young children (age 2) can be mediated by home stimulation (as measured by a home score). Rubio-Codina et al. (2016) show that the home score (based upon the stimulating environment of the household) mediates 38% of the socioeconomic gap in 6-42-month-old children. Galasso et al. (2017) study children in Madagascar, showing first that socioeconomic gaps are prevalent in all domains of child development in Madagascar, and second, that a strong home environment can close the gap by a magnitude of 12%-18%. Lastly, Lopez-Boo (2016) shows several different mediators of cognitive skill at ages 5 and 8 including caregiver's education.

3.3 Data

3.3.1 Measurement of Psychosocial Competencies

The data used in this study comes from the younger cohort of the Young Lives study in Peru. The younger cohort children have been followed from when they were 6 to 18 months old, up to age 12-13 years old across 20 randomly selected sentinel (community) sites. Of the 4 completed survey rounds, the two most recent rounds have included a sibling survey, where the younger sibling of the index child (the child who is followed for all four rounds) is interviewed for their anthropometrics, cognitive and non-cognitive skills (only available for the most recent round). For the first three rounds of the survey, there are 2,052 children who were interviewed in each of the rounds. Attrition from Round 3 to Round 4 is 7.3%, leaving 1,902 children (Azubuike and Briones 2016). Children who do not have full responses to their psychosocial statements, both in terms of their caregivers and their own, are removed. Thus, the sample is pared down to 503 sibling pairs. Peru is an important case study on socioeconomic inequalities as it has a high level of inequality. With a Gini index of 0.44, Peru is the 44th most unequal country in the world and 15th amongst Latin American countries per World Bank estimates (World Bank 2013). The inequality has been trending downwards over the last four years, highlighting the efforts of the government to improve conditions for low-income households.

The questionnaires contain statements to measure parent's, children's, and their sibling's, psychosocial competencies. Mothers are asked their level of agreement on four[13] different statements for pride, but not self-esteem. For children, pride is comprised of six[14] statements which focus on children's view of their things and achievements. It is a much more specific view of a child's feeling about their environment. Self-esteem is measured by the Rosenberg (1965) Self-Esteem scale which is comprised of eight questions.[15] The correlation between these two measures is 0.5. Even with this similarity, the pride index is used to test how good of a proxy the measure is of self-esteem, and maintains consistency with the other studies in the Young Lives research base examining these competencies.

Index children and their siblings are asked eight statements about their relationships with parents and peers in school and their community.[16] The measures of pride are set on a five-point Likert scale, while the self-esteem and relationship measures are a four-

[13]"I am proud of my clothes", "I feel proud of the job done by the household head", "The job I do makes me feel proud", and "I feel proud of my children"

[14]"I am proud of my clothes", "I am never embarrassed because I do not have the right books, pencils or other equipment for school", "I am proud of my achievements at school", "I am proud by/ashamed of the work I have to do", "I am proud of my shoes", and "I am proud that I have the correct uniform".

[15]"I'm as good as most other people", "Overall, I have a lot to be proud of", "In general, I like being the way I am", "I can do things as well as most people", "A lot of things about me are good", "I do lots of important things", "Other people think I am a good person", "When I do something, I do it well".

[16]Parents: "I like my parents", "My parents like me", "My parents and I spend a lot of time together", "I get along well with my parents", "My parents understand me", "If I have children of my own, I want to bring them up like my parents raised me", "My parents are easy to talk to", "My parents and I have a lot of fun together". Peers: "I make friends easily", "I get along with other kids easily", "I am popular with kids my own age", "Other kids want me to be their friend", "Most other kids like me", "Other kids want me to be their friend", "I have lots of friends", "I am easy to like".

point scale.[17] Each measure is standardised mean 0 and variance 1, summed together, and averaged to make a standardised index for each measure. This aims to approximate a measure which is associated with higher values of the included psychosocial competencies (Dercon and Krishnan 2009). Cronbach's Alpha, which measures the interrelatedness of the scales for the psychosocial competencies (Cronbach 1951), is calculated to measure the internal validity of these statements.

The formula is as follows:

$$\alpha = \frac{K_{\bar{c}}}{(\bar{v} + (K-1)\bar{c})} \tag{3.1}$$

The alpha measures how much of all the included statements are related to a specific concept or idea. A valid Cronbach's alpha generally is above 0.7. Here, the alpha's range from 0.55 to 0.82, so the measures are considered reliable.[18] The lowest measure is associated with maternal pride. Dercon and Krishnan (2009) state a few reasons for why this could be the case. For one, there could be a lack of understanding of the value of these concepts in the Peruvian context. In another sense, the measures are multidimensional in nature as the measure many concepts rather than a specific one. The index score of all the psychosocial statements lessens the concern somewhat as it captures the possible multidimensionality much better than each individual measure.

3.3.2 Descriptive Statistics

Descriptive statistics for wealth, household, and individual controls are found in Table 3.1. Household controls include a wealth index, size of the household, locality, caregiver's education and if the household has experienced an exogenous shock of any kind. The wealth index is the weighted average of consumer durables in the household, household services, and household quality. Individual controls include the age of the child, birth order, gender, standardised score on the PPVT test, if the child has

[17] These measures are adapted to the five-point scale for the analysis.

[18] The alpha for self-esteem is 0.69 and for pride is 0.66.

had a health problem since the last visit, and their height as an indicator of health. The sample is broken into older and younger siblings to present age differences.

Mothers and children have high levels of pride in the sample and good parental relationships, with the average response to the statements being "agree". Peer relationships and self-esteem are not as high however. Looking across the older and younger siblings, these results are driven by the older siblings having worse responses compared to their younger siblings. Half of caregivers in the sample have only a primary education, while the other half have either a secondary or tertiary level of education. 37% of the households are located in rural areas. PPVT scores are presented in raw form for ease of understanding. Older siblings perform better as expected, since it is a test of how many words you know.

In addition to Table 3.1, Table 3.2 provides these statistics for each quintile of wealth in the sample, per round. In addition, it is tested if there are significant differences between the upper, middle, and lower quintiles of wealth. Poorer children tend to have much less pride across the quintiles of wealth. Richer children have better relationships with their parents and higher self-esteem than those in the first quintile, but not at the median. Poorer households are larger, more likely to be in a rural area and have a caregiver with only a primary education. This encapsulates the disparity between rich and poor households; richer households are smaller, more educated, and more connected.

Table 3.1: Descriptive Statistics By Round

	(1) Sibling Sample		(2) Older Siblings		(3) Younger Siblings	
	mean	sd	mean	sd	mean	sd
Child's Pride, Raw Score	3.97	0.52	3.98	0.50	3.95	0.54
Maternal Pride, Raw Score	4.04	0.47	4.04	0.47	4.04	0.47
Rosenberg Self-Esteem Raw Score	3.50	0.56	3.14	0.33	3.86	0.50
Parent Relationship Raw Score	4.06	0.56	4.04	0.55	4.08	0.58
Friend Relationship Raw Score	3.58	0.55	3.52	0.53	3.63	0.56
Age in Years	10.43	1.76	11.95	0.22	8.90	1.23
Female	0.52	0.50	0.50	0.50	0.54	0.50
Household size	5.94	1.79				
Rural	0.37	0.48				
Caregiver has completed up to primary education	0.59	0.49				
Caregiver has completed up to the secondary education	0.08	0.27				
Caregiver has completed up to tertiary education	0.34	0.47				
Shock	0.68	0.47				
Raw PPVT Score	74.31	19.38	82.51	17.45	66.10	17.68
Child's Height (cm)	133.36	11.51	141.48	7.57	125.24	8.69

Table 3.2: Descriptives By Quintile of Wealth

	Summary Statistics					Difference In Means		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Q5	Q4	Q3	Q2	Q1	Q5 v Q1	Q5 v Q3	Q3 v Q1
Child's Pride Raw Score	4.13	4.05	3.96	3.92	3.90	0.23***	0.17**	0.07
	(0.53)	(0.54)	(0.50)	(0.50)	(0.52)	(0.06)	(0.06)	(0.05)
Maternal Pride Raw Score	4.08	4.06	4.10	4.03	3.99	0.08	-0.03	0.11*
	(0.44)	(0.47)	(0.49)	(0.46)	(0.46)	(0.05)	(0.05)	(0.04)
Rosenberg Self-Esteem Raw Score	3.60	3.54	3.51	3.46	3.47	0.14*	0.10	0.04
	(0.57)	(0.59)	(0.56)	(0.55)	(0.53)	(0.06)	(0.06)	(0.05)
Parent Relationship Raw Score	4.17	4.12	4.08	4.01	4.01	0.15**	0.09	0.06
	(0.51)	(0.58)	(0.58)	(0.54)	(0.58)	(0.06)	(0.06)	(0.05)
Friend Relationship Raw Score	3.65	3.61	3.62	3.52	3.55	0.10	0.03	0.07
	(0.54)	(0.58)	(0.52)	(0.54)	(0.57)	(0.06)	(0.06)	(0.05)
Household size	5.40	5.75	5.73	5.88	6.47	-1.07***	-0.34	-0.74***
	(1.54)	(2.01)	(1.87)	(1.53)	(1.81)	(0.17)	(0.19)	(0.17)
Rural	0.03	0.04	0.09	0.47	0.78	-0.75***	-0.06*	-0.69***
	(0.18)	(0.20)	(0.29)	(0.50)	(0.41)	(0.03)	(0.03)	(0.03)
Caregiver has completed up to primary education	0.21	0.26	0.49	0.74	0.85	-0.64***	-0.29***	-0.36***
	(0.41)	(0.44)	(0.50)	(0.44)	(0.36)	(0.04)	(0.05)	(0.04)
Caregiver has completed up to the secondary education	0.13	0.05	0.05	0.10	0.06	0.06	0.08*	-0.01
	(0.33)	(0.23)	(0.22)	(0.30)	(0.24)	(0.03)	(0.03)	(0.02)
Caregiver has completed up to tertiary education	0.67	0.68	0.45	0.16	0.09	0.58***	0.21***	0.37***
	(0.47)	(0.47)	(0.50)	(0.37)	(0.28)	(0.05)	(0.06)	(0.04)
Shock	0.62	0.63	0.65	0.79	0.66	-0.04	-0.03	-0.01
	(0.49)	(0.48)	(0.48)	(0.41)	(0.47)	(0.05)	(0.06)	(0.04)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Groups divided by quintiles of wealth

3.3.3 Non-Parametric Analysis

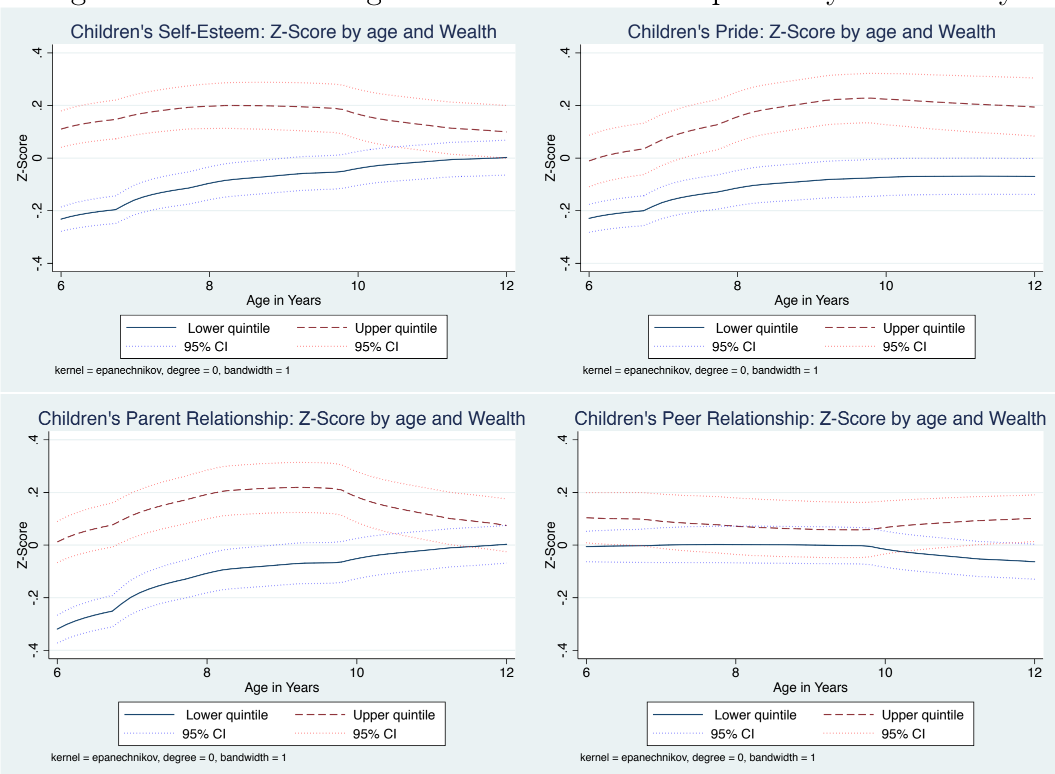
Non-parametric analysis can be a useful tool to illustrate how skills fluctuate over a given period. The siblings component of the study means that children aged 6 to 12 years old are studied at once. Figure 3.1 shows the three main variables in this analysis. Each outcome measure is standardised by age in years and used in the function. Using the Epanechnikov kernel function, smoothed graphs of pride and agency constructed by wealth quintiles can be created. Equation 3.2 is the smoothing function.

$$\hat{Y}(X) = \frac{\sum_{i=1}^{N} K_{h\lambda}(X_O, X_i) Y(X_i)}{\sum_{i=1}^{N} K_{h\lambda}(X_O, X_i)} \tag{3.2}$$

where the kernel function, K, is the Epanechnikov function:

$$K_{h\lambda} = \frac{3}{4} \frac{(1 - \frac{u^2}{5})}{\sqrt{5}} \tag{3.3}$$

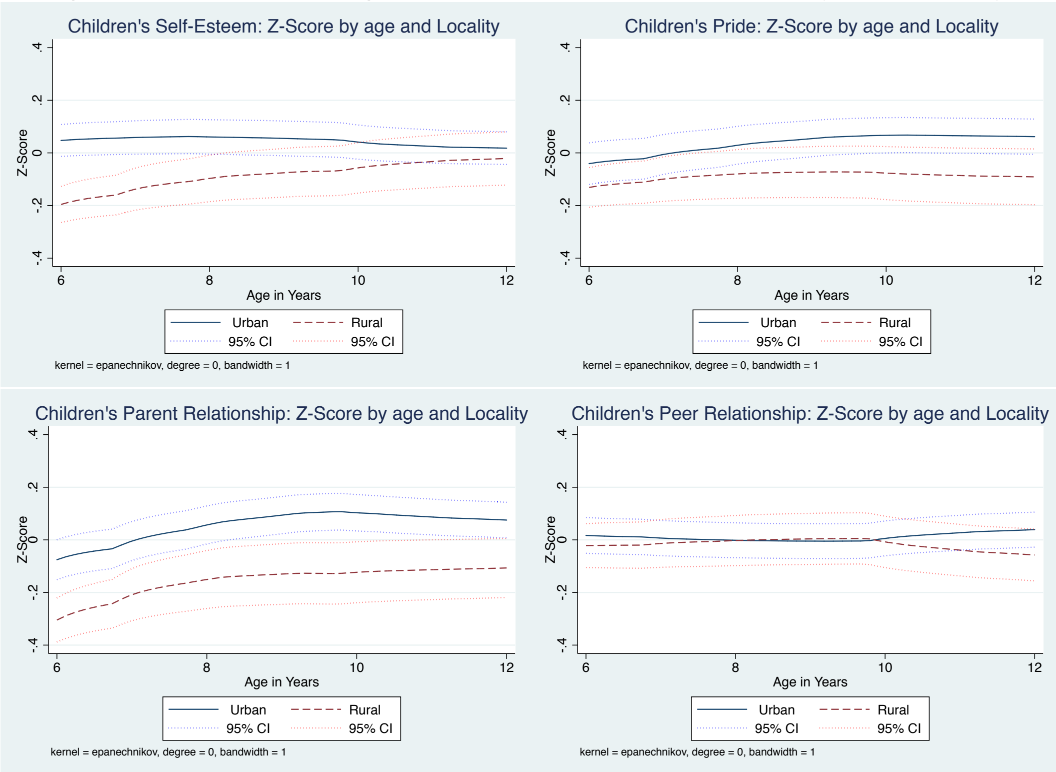
Figure 3.1: Smoothed figures for outcome and explanatory variables by wealth



For self-esteem, the gaps are formed before age 6 and persist up to around age 9, where they flatten and close. The parent-child relationship follows a similar trend, showing that as children are exposed to more they become less dependent on their wealth status. Peer relationships are stable and indifferent up to age 10, where they begin to diverge slightly. Lastly, gaps in pride increase up to a certain point and stabilise. This shows that pride, in contrast to self-esteem and the relationship measures, is more related to wealth and stabilises once children are firmly entrenched in schooling.

Figure 3.2 reproduces the same graphs using locality rather than wealth as the differentiator. The figures follow the same pattern as the above except in terms of the parental relationship measure, which is significantly different throughout the age profile. The result is possibly due to the economic situation of a rural household. Children have more responsibility working, and as such, likely relate more to their parents in this sense rather than a familial sense.

Figure 3.2: Smoothed figures for outcome and explanatory variables by locality



3.4 Empirical Specification

The above analysis has illustrated the presence of significant gaps in children's self-esteem and pride between different levels of socioeconomic status. The next step is to parameterise and estimate the wealth gradient for the sample of children. A recentered influence function (RIF) enriches the analysis by studying the behaviour of determinants across the distribution of children's self-esteem and pride. The sibling's component of the survey can be exploited to obtain more robust estimates of the impact of relationships on the development of these constructs. Finally, decomposing the gaps by socioeconomic status and locality can estimate how equalising inputs can close psychosocial gaps between rich and poor children.

3.4.1 Capturing Wealth Gradients and Determinants

A production function for skills (Todd and Wolpin 2003, 2007) is used to explain how children's psychosocial skills are developed between age 6 and 12. This specification builds on Dercon and Sanchez (2011) and is chosen because it treats psychosocial competencies as constructs that are produced through investment into factors. For example, Dercon and Sanchez (2011) identify malnutrition as a channel which socioeconomic status specifically travels through to affect psychosocial competency development. Equation 3.4 illustrates this production function, which is a simplification of what is presented in Dercon and Sanchez (2011).

$$\theta_{it}^{C} = f(\theta_{it}^{M}, r_{it}^{p}, r_{it}^{F}, X_{it}, \mu_i) \tag{3.4}$$

$$\theta_{it}^{C} = \alpha + \sum_{i=1}^{4} \beta_j WQ_{ij} + \beta_5\theta_{it}^{M} + \beta_6 r_{it}^{P} + \beta_7 r_{it}^{F} + \beta_8 caredu + \beta_9 X_{it} + \mu_i \tag{3.5}$$

$$where\, C, P, F \,\epsilon\, Child,\ Mother,\ Friend$$

Here, θ_{it}^{C} is the child's pride or self-esteem. θ_{it}^{M} is the measure of maternal pride, a proxy of their self-esteem. For theoretical simplicity, inputs are contemporaneous.

Relationship, r, denotes the quality of the relationship between parents, P, and peers, F, respectively. Maternal pride is a potential channel for development as it indicates a component of the household environment and proxies for maternal self-esteem. A proud household should lead to prouder children, all things equal. Higher quality relationships (reflected by higher scores) are also expected lead to higher self-esteem and pride. As children interact with others positively, they would view themselves more positively and be prouder of who they are and what they have.

X_{it} includes the household and individual level controls such as caregiver's education, household size, gender, locality, whether the household experienced a shock, age, birth order of the child, and the child's height and any health problems experienced. Also, contained within X_{it} is the child's contemporaneous PPVT score. The addition of the contemporaneous cognitive score acts as a control for the rank of the intelligence of the child amongst its peers in the study. Lastly, μ_i is the error term which includes the genetic endowment that a child has from birth as well as unobservable factors such as the household dynamics not captured by birth order. This is the main identification concern.

Equation 3.5 linearises the production function in Equation 3.4. Caregiver's education is extracted from X_{it} because it is an indicator of the parent's level of human capital, the intergenerational aspect of socioeconomic status in the specific household, and the level of support they may have received as a child from their parents and caregivers. Higher levels of education suggest that parents had parents who invested in their development.

As self-esteem and pride are constructs which capture how individuals view their place in the world, examining mean effects misses out on how children on either end of the spectrum differ. It is expected that wealth gradients and determinants of self-esteem and pride would vary accordingly in these children, as they grow up in a different environment and derive their pride and agency from different sources. Firpo, Fortin and Lemieux (2009) propose an unconditional quantile regression method to estimate relationships across different points of the distribution. Normally used for wage distributions, the method can be used in this literature to describe heterogeneity across the distribution of skills.

The method generates a recentered influence function of a specific quantile (or variance or Gini coefficient) and uses it as the dependent variable in a standard OLS regression. The influence function, $IF(Y, Q_t) = (\frac{(\tau - \mathbb{1}\{y \leq Q_\tau\})}{(f_y(Q_\tau))}$, of a distributional statistic describes the impact of a set of covariates on a specific distributional measure. The indicator function $\mathbb{1}\{y > Q_\tau\}$ denotes whether a certain outcome is above or below a given quantile in the distribution. Recentering this measure allows it to be aggregated back to statistics of interest (Fortin, Lemieux, Firpo 2009, p.73). The RIF is empirically created by first computing sample quantiles and estimating kernel densities. These densities are created individually for each group being studied; in this case poor and non-poor and urban and rural.

$$RIF(y; Q_\tau) = Q_\tau + \frac{(\tau - \mathbb{1}\{y \leq Q_\tau\})}{(f_y(Q_\tau))} = c_{1,\tau} \cdot \mathbb{1}\{y > Q_\tau\} + c_{2,\tau} \tag{3.6}$$

3.4.2 Sibling Differences

The results from the standard OLS specification are consistent if the unobserved factors (such as many components of the household environment) are orthogonal to the included inputs (Todd and Wolpin 2003, 2007). A household fixed effects model can improve identification by removing some of these unobserved factors, such as the quality of inputs in the household and household wealth effects by differencing across siblings. The Young Lives survey provides an opportunity to do this through information on the index children's younger siblings. These younger siblings answer the same pride, Rosenberg Self-Esteem and parental and peer relationship statements at age 7 and 8 as the index children, making it possible to use a sibling difference estimator and remove these unobservable factors.

The sibling difference estimator is easy to implement. Equation 3.7 is the empirical specification for each sibling with sibling variant inputs.

$$(\theta^C_{ij} - \theta^C_{ij^*}) = \beta_1(r^P_{ij} - r^P_{ij^*}) + \beta_2(r^F_{ij} - r^F_{ij^*}) + \beta_3(X_{ij} - X_{ij^*}) + (\mu_{ij} - \mu_{ij^*}) \tag{3.7}$$

Here, r^P_{ij} and r^F_{ij} measure the quality of relationships with parents and friends respectively. Because the statements are answered in the same interview round, there is no difference in reported maternal pride or education for both children. As a result, the only measured variation in parental inputs is the self-reported relationship scores. One weakness of the sibling's survey is that these children are the next children in line in the household, so it is harder to capture heterogeneity throughout the household sample.

The estimates of β_1, the effect of the quality of relationship with parents, and β_2, the effect of the quality of relationship with friends, are consistent if the explanatory variables are exogenous to the unobserved factors for the index child and their younger sibling. I argue that this is the case for relationships with friends, because any unobservable factors (such as household environment or structure for example) should be unrelated with how your friends relate with you. In the case of parental relationship, the biggest factors of relationship quality would be factors such as gender, age, birth order and revealed intelligence, all of which are controlled for in the above specification. The main concerns beyond this would be whether individual children are treated different for reasons beyond these. With the set of controls, these concerns are lessened, and the presented effects can be considered as causal. At the very least, they are the next best alternative to causal, in absence of a viable instrument.

3.4.3 Identifying the Role of Policy

Decomposing the self-esteem and pride gaps between rich and poor children can be done through an Oaxaca (1973)-Blinder (1973) decomposition. The decomposition can be advanced across quantiles using the Firpo, Fortin and Lemieux (2009) procedure to provide a more in-depth picture of potential policy impacts. Equation 3.8

applies the decomposition to this model. Children are separated into two groups either by socioeconomic status or locality. In the case of socioeconomic status, there is a poor group (p), which are those in the first quintile of the wealth index and a non-poor group (np), which are all the other children. The first component of the equation is the component explained by different levels of the variables in the model, denoted by the difference between the mean covariates of each group multiplied by the coefficients of the poor group. The second component is the unexplained component, denoted by the difference in coefficients for each group multiplied by the mean of the covariates in the non-poor group.

$$\bar{D}_p - \bar{D}_{np} = [f(\bar{X}_p\hat{\beta}_p) - f(\bar{X}_{np}\hat{\beta}_p)] + [f(\bar{X}_{np}\hat{\beta}_p) - f(\bar{X}_{np}\hat{\beta}_{np})] \tag{3.8}$$

3.5 Results

3.5.1 Determinants of self-esteem and pride

Table 3.3 examines the associations between the explanatory variables and self-esteem and pride first for the entire sample, and then subsamples for older and younger children. The results estimate Equation 3.5, and show that the wealth gradients are non-existent across the sample, suggesting that any wealth effects are now stable.

Table 3.3: Pooled Child's Psychosocial Competencies

	Pooled	Index Child	Younger Sibling
Panel A: Self-Esteem			
Top Quintile of Wealth	0.061	0.031	0.068
	[-0.03,0.2]	[-0.1,0.2]	[-0.1,0.3]
Maternal Pride	-0.024	-0.022	-0.024
	[-0.06,0.02]	[-0.08,0.03]	[-0.07,0.02]
Parent Relationship Score	0.33***	0.33***	0.31***
	[0.2,0.4]	[0.2,0.5]	[0.2,0.4]
Peer Relationship Score	0.40***	0.38***	0.43***
	[0.3,0.5]	[0.3,0.5]	[0.3,0.5]
Caregiver's Education	0.0017	-0.0026	0.0054
	[-0.008,0.01]	[-0.01,0.009]	[-0.008,0.02]
Panel B: Pride			
Top Quintile of wealth	0.045	0.12	-0.014
	[-0.06,0.2]	[-0.06,0.3]	[-0.2,0.2]
Caregiver Self-Esteem	0.013	0.015	0.010
	[-0.03,0.05]	[-0.04,0.07]	[-0.06,0.08]
Parent Relationship Score	0.33***	0.37***	0.30***
	[0.3,0.4]	[0.3,0.5]	[0.2,0.4]
Peer Relationship Score	0.27***	0.24***	0.29***
	[0.2,0.3]	[0.1,0.4]	[0.2,0.4]
Caregiver's level of education	0.012***	0.0073	0.016***
	[0.004,0.02]	[-0.003,0.02]	[0.005,0.03]
Observations	1006	503	503

95% confidence intervals in brackets. Robust standard errors clustered at community level.
Wild Bootstrap hypothesis testing confirms the results considering small cluster amount.
Controls include gender, household size, age in years, birth order, locality,
standardised PPVT score, child's height, whether the house experienced a shock, and region fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The significant determinants of self-esteem and pride are the relationship measures and caregiver education (for pride). For self-esteem the largest predictor is peer relationships, while for pride it is the parent-child relationship. In either case, children with strong relationships with those in their surroundings receive positive feedback from their interactions which boost the way they view themselves. As they begin to work, this feedback enables them to be successful.

Table 3.4 estimates Equation 3.6 and examines the results across different quantiles of children's self-esteem and pride. The 10th percentile being children with low stocks of these competencies, the 50th percentile is the median, and the 90th percentile being those with high stocks. OLS results are reproduced from above for convenience in Column 1.

Table 3.4: Pooled Determinants Across Quantiles

	OLS	10th Percentile	50 Percentile	90th Percentile
Panel A: Self-Esteem				
Caregiver's Education	0.0017	0.014	-0.0030	0.0071
	[-0.008,0.01]	[-0.004,0.03]	[-0.01,0.008]	[-0.01,0.03]
Maternal Pride	-0.024	-0.035	-0.019	-0.0024
	[-0.06,0.02]	[-0.2,0.08]	[-0.07,0.04]	[-0.1,0.10]
Parent Relationship Score	0.33***	0.22***	0.37***	0.44***
	[0.2,0.4]	[0.1,0.3]	[0.3,0.5]	[0.3,0.6]
Peer Relationship Score	0.40***	0.45***	0.33***	0.52***
	[0.3,0.5]	[0.3,0.6]	[0.2,0.4]	[0.3,0.7]
Caregiver's Education	0.0017	0.014	-0.0030	0.0071
	[-0.008,0.01]	[-0.004,0.03]	[-0.01,0.008]	[-0.01,0.03]
Panel B: Pride				
Top Quintile of Wealth	0.045	-0.096	0.061	0.32**
	[-0.06,0.2]	[-0.4,0.2]	[-0.1,0.2]	[0.03,0.6]
Maternal Pride	0.013	0.033	-0.0038	0.032
	[-0.03,0.05]	[-0.08,0.1]	[-0.07,0.06]	[-0.05,0.1]
Parent Relationship Score	0.33***	0.31***	0.25***	0.39***
	[0.3,0.4]	[0.1,0.5]	[0.2,0.3]	[0.3,0.5]
Peer Relationship Score	0.27***	0.25***	0.17***	0.37***
	[0.2,0.3]	[0.09,0.4]	[0.10,0.2]	[0.2,0.5]
Caregiver's Education	0.012***	0.017	0.009*	0.0004
	[0.004,0.02]	[-0.005,0.04]	[-0.002,0.02]	[-0.02,0.02]
Observations	1006	1006	1006	1006

95% confidence intervals in brackets. Robust standard errors clustered at community level.
Wild Bootstrap hypothesis testing confirms the results considering small cluster amount.
Controls include gender, household size, age in years, birth order, locality,
standardized PPVT score, child's height, whether the house experienced a shock, and region fixed effects.
Distributional effects computed through RIF-procedure by Firpo, Fortin and Lemieux (2009).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In terms of self-esteem, children in the high and low part of the distribution benefit the most from their relationships with their peers. At the top end of the scale, this could be a reinforcing effect. Richer children have a certain view of the world that is reinforced by their peers. On the bottom end, children look to their relationships as a supporting tool and gain more from those relationships than those in the median. For the relationship with parents, it is an increasing trend, providing more evidence of the reinforcing effect as children have higher levels of self-esteem. In terms of pride, there is a similar U-shape relationship across the quantiles in both relationship indicators.

An additional table in the appendix examines subsamples for brothers and sisters and for urban and rural localities. The results, while showing some specific points of heterogeneity, are generally the same. Interaction specifications are also considered, in terms of being the older or younger sibling and age, and are available on request.

3.5.2 Sibling Differences

Table 3.5 exploits the siblings component of the survey for more robust estimates of the effect of relationships on self-esteem and pride. Maternal pride cannot be used in this specification as it is the same for both the index child and their younger sibling. The results of the full sample are reproduced for comparison in Column 1.

The siblings estimates for self-esteem and pride are similar to the pooled specification, showing that the relationship is strong. The controls included in the Young Lives data are thorough enough to capture the unobserved household effects. There are large positive effects of good relationships with peers and parents of nearly 0.4 SD, with the peer relationship being slightly larger than parents. This likely reflects the fact that children are in school for long periods of the day, and their peers are more important drivers of how they view their place in the world.

Table 3.6 presents subsample analysis by type of sibling pair and locality. Sister pairs are driving the magnitude of the effect in both competencies, while peers are more important for self-esteem in rural settings. Notably in the brother-sister pairs, being female is a positive determinant for self-esteem, but not pride. Overall though, the results are consistent with those seen throughout this analysis.

Table 3.5: Siblings Difference

	Pooled	Sibling Difference
Panel A: Self-Esteem		
Top Quintile of Wealth	0.061	
	[-0.03,0.2]	
Parent Relationship Score	0.33***	0.36***
	[0.2,0.4]	[0.3,0.5]
Peer Relationship Score	0.40***	0.39***
	[0.3,0.5]	[0.3,0.5]
Panel B: Pride		
Top Quintile of Wealth	0.045	
	[-0.06,0.2]	
Parent Relationship Score	0.33***	0.29***
	[0.3,0.4]	[0.2,0.4]
Peer Relationship Score	0.27***	0.34***
	[0.2,0.3]	[0.3,0.4]
Observations	1006	1006

Robust standard errors clustered at community level.
Wild Bootstrap hypothesis testing confirms the results considering small cluster amount.
Controls include gender, age in years, locality, standardized PPVT score, and child's height.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.5.3 Closing Socioeconomic and Locality Gaps

From a policy perspective, it is important to determine how gaps between rich and poor, and urban and rural children could be closed by boosting maternal pride and promoting stronger relationships between children and their parents and peers. By equalising inputs in rich and poor children, or by locality, it can be estimated how much of this gap closes at the mean, and across the distribution. Most of the previous literature focuses on interventions implemented in the early childhood (Heckman et al. 2010). Kautz et al. (2014) summarises the success of a few adolescent programmes in North America, finding some benefits in the way of improved cognitive ability and decreased likelihoods of violence (Cook et al. 2014; Oreopoulos et al. 2017; Tierney et al. 1995). Additionally, the after-school arts and sports programme evaluated by

Table 3.6: Sibling Difference Results by Subsamples

	Brothers	Sisters	Differing Pairs	Urban	Rural
Panel A: Self-Esteem					
Parent Relationship Score	0.37***	0.46***	0.33***	0.37***	0.37***
	[0.2,0.6]	[0.3,0.6]	[0.2,0.5]	[0.3,0.5]	[0.2,0.5]
Peer Relationship Score	0.42***	0.47***	0.33***	0.37***	0.45***
	[0.3,0.6]	[0.3,0.7]	[0.1,0.5]	[0.2,0.5]	[0.3,0.6]
Panel B: Pride					
Parent Relationship Score	0.25**	0.31***	0.30***	0.31***	0.26***
	[0.03,0.5]	[0.1,0.5]	[0.2,0.4]	[0.2,0.4]	[0.1,0.4]
Peer Relationship Score	0.26**	0.53***	0.28***	0.34***	0.34***
	[0.04,0.5]	[0.4,0.7]	[0.2,0.4]	[0.2,0.5]	[0.3,0.4]
Observations	232	276	498	636	370

95% confidence intervals in brackets. Robust standard errors clustered at community level.
Wild Bootstrap hypothesis testing confirms the results considering small cluster amount.
Controls include gender, age in years, locality, standardised PPVT score, and child's height.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Krishnan and Krutikova (2013) in India shows positive improvements in children's self-esteem. Examining effects across the distribution then provides more clarity to who benefits the most from equalising inputs. Tables 3.7 and 3.8 show these results for socioeconomic status and locality respectively, estimating Equation 3.8. Only the explained components are included in this table for brevity.[19]

[19] Full results available on request.

Table 3.7: Quantile Decomposition by Wealth

	Oaxaca-Blinder	10th Percentile	50th Percentile	90th Percentile
Panel A: Self-Esteem				
Total Explained	0.051	0.08	0.108	0.17
	[-0.026,0.13]	[-0.12, 0.26]	[0.031, 0.22]	[0.028, 0.34]
Explained Component				
Maternal Pride	-0.002	-0.002	-0.0034	0.005
	[-0.010,0.0059]	[-0.022, 0.019]	[-0.019, 0.006]	[-0.007, 0.031]
Parent Relationship Score	0.026	0.018	0.027	0.034
	[-0.0055,0.058]	[-0.019, 0.069]	[-0.015, 0.068]	[-0.03, 0.09]
Peer Relationship Score	0.019	0.021	0.018	0.026
	[-0.017,0.055]	[-0.025, 0.089]	[-0.026, 0.066]	[-0.04, 0.12]
Caregiver's Education	0.0027	0.036	-0.022	0.035
	[-0.034,0.039]	[-0.054, 0.13]	[-0.08, 0.05]	[-0.06, 0.16]
Total Explained	0.051	0.073	0.058	0.082
	[-0.026,0.13]	[-0.075, 0.25]	[-0.058, 0.18]	[-0.09, 0.35]
Panel B: Pride				
Total Difference	0.12**	0.12	0.055	0.17
	[0.026,0.21]	[-0.08, 0.32]	[-0.053, 0.193]	[-0.046, 0.37]
Explained Component				
Maternal Pride	-0.002	0.003	0.001	0.006
	[-0.010,0.0059]	[-0.028, 0.041]	[-0.0144, 0.013]	[-0.01, 0.026]
Parent Relationship Score	0.026	0.02	0.02	0.026
	[-0.0055,0.058]	[-0.013, 0.061]	[-0.017, 0.054]	[-0.021, 0.074]
Peer Relationship Score	0.019	0.013	0.01	0.021
	[-0.017,0.055]	[-0.015, 0.052]	[-.012, 0.044]	[-0.025, 0.083]
Caregiver's Education	0.0027	0.034	0.03	0.031
	[-0.034,0.039]	[-0.101, 0.173]	[-0.03, 0.087]	[-0.083, 0.135]
Observations	1006	1006	1006	1006

Standard errors in parentheses. Households declared poor if in the lower quintile of wealth.
Differences in added coefficients and total cells due to rounding.
A bootstrap of the entire process is performed over 100 reps to obtain standard errors.
Controls include gender, household size, age in years, birth order, locality, standardized PPVT score, child's height, whether the house experienced a shock, and region fixed effects.
Significance stars are not included for the distribution effects.
Results for unexplained components removed for brevity, but available on request.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

It is apparent that there are gaps between rich and poor groups across the mean and distribution for pride, which are estimated to be nearly closed by equalising inputs. This is not a hugely surprising result, as the questions are aimed at capturing the wealth component of pride (i.e. are you proud of your things?). Equalising the relationship with parents closes approximately 21% of the gap at the mean, an outcome which supports the earlier results. The effects are less precise in the quantiles, suggesting, that any improvements in policy will only affect those around the mean, and not at the ends of the distribution. For self-esteem, there is very little difference between rich and poor households.

Table 3.8 reproduces the above table using locality rather than socioeconomic status as the differentiator between households. Gaps between urban and rural households are similar in magnitude to socioeconomic status, but there is a large disparity in the way children relate to parents in rural households. Improving inputs in these households can close 80% of the socioeconomic gap, while closing 31% of the gap in children's pride. This is an important result, because it shows that the conditions which children grow up in are much different in urban and rural areas and that it is an indirect factor in forming gaps in children's self-esteem and pride. It also highlights the differences between the two measures; the generalised measure of self-esteem is more relational, while pride in this setting is based more on wealth.

Table 3.8: Quantile Decomposition by Locality

	Oaxaca-Blinder	10th Percentile	50th Percentile	90th Percentile
Panel A: Self-Esteem				
Total Difference	0.088**	0.023	0.12	0.1
	[0.031,0.15]	[-0.20, 0.20]	[0.004, 0.22]	[-0.18, 0.27]
Explained Component				
Maternal Pride	-0.002	-0.0003	-0.005	0.01
	[-0.008,0.005]	[-0.023, 0.026]	[-0.02, 0.01]	[-0.006, 0.04]
Parent Relationship Score	0.074***	0.06	0.083	0.085
	[0.049,0.098]	[0.011, 0.11]	[0.025, 0.16]	[0.026, 0.14]
Peer Relationship Score	0.015	0.017	0.013	0.018
	[-0.012,0.043]	[-0.03, 0.08]	[-0.02, 0.06]	[-0.027, 0.078]
Caregiver's Education	-0.0037	0.025	-0.06	0.043
	[-0.035,0.027]	[-0.078, 0.12	[-0.15, 0.025]	[-0.1, 0.17]
Whether HH is Poor	-0.00083	-0.011	-0.028	0.034
	[-0.052,0.050]	[-0.15, 0.14]	[-0.16, 0.10]	[-0.12, 0.17]
Total Explained	0.095**	0.12	0.025	0.22
	[0.028,0.16]	[-0.16, 0.35]	[-0.124, 0.144]	[-0.062, 0.49]
Panel B: Pride				
Total Difference	0.14***	0.11	0.076	0.19
	[0.074,0.21]	[-0.11, 0.33]	[-0.026, 0.21]	[-0.006, 0.37]
Explained Component				
Maternal Pride	0.005	0.001	0.003	0.007
	[-0.003,0.013]	[-0.027, 0.04]	[-0.012, 0.01343]	[-0.013, 0.04]
Parent Relationship Score	0.075***	0.06	0.06	0.08
	[0.045,0.10]	[.014, 0.13]	[0.03, 0.1]	[0.034, 0.15]
Peer Relationship Score	0.0097	0.008	0.006	0.016
	[-0.0080,0.027]	[-0.02, 0.03]	[-0.02, 0.02]	[-0.043, 0.07]
Caregiver's Education	0.052**	0.04	0.05	0.077
	[0.012,0.092]	[-0.094,0.22]	[-0.011, 0.12]	[-0.015, 0.20]
Whether HH is Poor	0.053	0.026	0.04	0.076
	[-0.0061,0.11]	[-0.23, 0.34]	[-0.086, 0.16]	[-0.026, 0.24]
Total Explained	0.24***	0.26	0.20	0.27
	[0.17,0.32]	[0.017, 0.55]	[0.09, 0.34]	[0.04, 0.61]
Observations	1006	1006	1006	1006

Standard errors in parentheses. Households separated by location in a rural or urban area.
Differences in added coefficients and total cells due to rounding. Households declared poor if in the lower quintile of wealth.
A bootstrap of the entire process is performed over 100 reps to obtain standard errors.
Controls include gender, household size, age in years, birth order, locality, standardized PPVT score,
child's height, whether the house experienced a shock, and region fixed effects.
Significance stars are not included for the distribution effects.
Results for unexplained components removed for brevity, but available on request.
* $p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

3.6 Discussion and Conclusions

The previous literature has established that the first five years of life are critical for the development of a child's skills. However, it has not addressed how gaps which are formed in this period behave as time passes due to data constraints. Additionally, there has been little focus in the literature on the existence of causal links between parents, peers, and children's self-esteem and pride. The Young Lives study in Peru enables deeper research into these topics due to the longitudinal and sibling component of the data, as household unobserved impacts can be removed.

The main findings of the paper are as follows. First, self-esteem and pride are similar, but distinct processes. Pride, measured as a subset of self-esteem focused on things in a child's environment, is much more dependent on wealth, as it deals with their view of the worth of their things and achievements within it. Poor households, compared to rich households, will always have a worse view of things, such as materials and clothes, than richer households. Generalised self-esteem however does not behave the same way, as there is only a small gap between rich and poor households at the mean. Non-parametric analysis of the two psychosocial competencies show stable socioeconomic gaps for pride, while the self-esteem gaps close as children age. The results are similar when locality is considered, but do show that there is a significant and stable difference between urban and rural children's relationships with their parents.

The second finding of the paper is that while there are wealth gaps in children, these gaps are stable as children age out of early childhood. The main determinants at this point are children's relationship with peers and their parents and caregiver's education. The estimates are close for each outcome, with parent and peer relationships accounting for 50-80 percent of the variance. The closeness of two estimates reflects the similarity between the two measures, and supports the use of pride as a proxy of self-esteem in the past Young Lives literature. The OLS estimates yield two points of note; first, that self-esteem (and by extension pride) is a construct which reflects the child's environment and second, there is an intergenerational component. The intergenerational component comes from the fact that the way parents treat their children, likely carried on from their parents, is reflected in their own children. Deeper analysis across the distribution shows that children with high or low stocks of skills derive their skills differently, suggesting that the way children view the world is reflected in the way their self-esteem is developed. This finding is strengthened when the siblings component of the Young Lives survey is used. By differencing between siblings, any household unobservable inputs to children, provided they are equal to both siblings, are removed, improving the identification of the variables of interest. The estimates are comparable to the OLS results, confirming that relationships are an important determining factor for children's self-esteem and pride.

The last finding examines how socioeconomic and locality gaps would be closed by equalising inputs between these households. For socioeconomic status, the results are concentrated in the mean, but show that equalising relationships between children and their parents can lead to the socioeconomic gap in pride closing by 21%. In the case of self-esteem, the overall gaps are smaller than pride, and there are no statistically significant effects from equalising inputs. This supports the idea that pride is dependent on wealth, while self-esteem is a deeper measure. In terms of locality, relationships with parents have a significant role in closing pride gaps by 31% and self-esteem by nearly 80%. There is a huge difference in the way children relate to their parents by urban and rural location. Combined with the socioeconomic evidence, there is a policy argument to target parental relationships in poor, rural areas in Peru.

The paper is unique to the existing literature in that it captures a deeper connection between parents, peers and children with measurements on the quality of their relationships. The results show that maternal pride, an indicator of the household environment, has no significant effect on children's self-esteem and pride once the relationship with parents is considered. Compared to the existing literature which focuses on intergenerational transfers, these results suggest that the existing estimates are lower bound effects of parents on their children. The previous literature has presented correlations which vary from 0.11-0.24 SD depending on the age being studied (Anger 2012; Loehlin 2011; Duncan 2005). Early childhood results in Rubio-Codina et al. 2016 (an effect size of 0.11 SD) and Fernald et al. 2012 (an effect size of 0.12 SD). The parental relationship measure used in this analysis yields estimates of approximately 0.33SD, which is robust to the removal of unobserved household factors through the sibling's difference estimator. Altogether, this is strong evidence of an intergenerational component to the transfer of these skills.

In terms of closing socioeconomic gaps, a strong relationship with parents leads to socioeconomic gaps closing by 21% at the mean for pride. This is estimate is comparable to previous literature, namely Rubio-Codina et al. (2016) which repots mediation effects of 38% in Colombia and Fernald et al. (2012) which reports 20% in Peru. Galasso et al. (2017) simulates that equalising home environments in Madagascar would close skill gaps by 12%-17%, close to the 14%-20% estimated here for children's pride. These are modest gains in reducing socioeconomic inequality and they emphasise the importance of boosting relationships and communities to provide children the environment they need to develop fully. Closing locality gaps through improving relationships leads to much larger gains in both self-esteem and pride. To the author's knowledge, there are no other examples of examining psychosocial competencies by locality in such a deep manner; a unique contribution of this paper which can guide future policy.

The results here would benefit from two additional features which are not available in this sample. First, additional data looking at the prevalence of risky behaviour and economic outcomes would be useful to see how stocks of these competencies translate to later life behaviours and outcomes. Recent research by Favara and Sanchez (2017) from the Young Lives study shows a negative relationship between children's self-esteem and their risky behaviours without truly accounting for the experience of children in their most formative years. This comprehensive picture of development is essential for forming the most impactful policy. Second, data on twins would be useful to further improve the identification of the siblings difference estimator. An identification concern still exists in these estimates relating to unobserved genetic impacts. With twins data, that concern is reduced as this component is more easily assumed away, providing stronger results (although Bhalotra and Clarke [2016] provide a caveat to this).

Given the available data, the presented estimates are arguably the best possible identification of the effects of parent and peer relationships on children's self-esteem and pride. Improving relationships between peers and parents is a key factor to consider when examining how children's self-esteem and self-efficacy are developed. The modest evidence of closing socioeconomic gaps through equalising the quality of parental relationships between parents and their children highlights an important channel which policymakers can use to improve children's chance of success in the future. This is especially true in rural areas compared to urban areas, where there is a large difference in the way parents and children relate. The nature of living in rural households is different, where parents are reliant on older children to work and earn income for their households. Policy which aims to foster better relationships, through community activities and institutions, can lead to more proud and high self-esteem children, and help households grow out of poverty. It is important that these policies account for the unique nature of locality, and not neglect the income generating activities of the household. This type of intervention would be a useful compliment to existing anti-poverty measures in poor, rural areas.

4 Exploring Early Childhood Development Effects from a Honduran Cash Transfer targeted at Older Siblings

with Florencia Lopez-Boo[20]

4.1 Introduction

Conditional cash transfer programmes (CCT's) have been used in many settings to reduce poverty and improve human capital and health. Low-income households have credit and income constraints which increase the incentive for households to remove their older children from school and into income generating activities. A poverty trap persists, as children lack the skills to be competitive in the labour market, and are stuck in low-income and risky jobs. The overall goal of conditional cash transfers is to lower the opportunity cost between human capital promoting activities (such as improved school enrolment and attendance) and labour by providing cash to household's conditional on the household sending a child to school for longer or by utilising health services. Children with more human capital are more competitive and successful in the labour market, increasing the probability of their households growing out of poverty. In many cases, evaluations of these CCTs, often based on random assignment, focus directly on the impacts on the main beneficiaries of the programme (i.e. the children who are induced to attend school/health check-ups through the conditionality) to determine the success of the transfer (Fiszbein and Schady 2009), rather than measuring effects that extend beyond the stated programme goals.

While there is mixed evidence in the literature on the indirect effects of CCTs on non-targeted children's nutritional status, (Fiszbein and Schady 2009; Lagarde, Haines, and Palmer 2009; Araujo, Bosch and Schady 2016), less is known about the presence of

[20]We deeply thank Catherine Porter and Pablo Ibarraran for their comments, as well as Felipe Sarmiento for earlier research assistance.

these effects on other domains of child development. Two randomised evaluations from Ecuador and Nicaragua report estimates of the impact of cash transfer programmes on child cognitive and language development. In Ecuador, the Bono de Desarrollo Humano (BDH) had a significant impact on cognitive and behavioural outcomes among children 36-59 months in the poorest households, with an effect size of 0.18 standard deviations (Paxson and Schady 2010). For younger children treated at 12-35 months, the treatment resulted in more words being spoken at follow-up (Fernald and Hidrobo 2011). In Nicaragua, the Atención a Crisis pilot programme improved the cognitive, language, and behavioural development of children 0–5 years of age by 0.12 standard deviations (Macours, Schady, and Vakis 2012). The evaluation design in Nicaragua also allows for an analysis of the effects of bigger and smaller transfers. Comparisons between the two treatment groups show that overall consumption increased by more in the group that received the larger transfers, as expected. However, child development outcomes did not improve by more in this group, suggesting that something other than (or in addition to) the cash was at work. The Atención a Crisis programme also changed various behaviours that are associated with better child outcomes (for example, parents were more likely to tell stories, sing, or read to their children). The changes in these behaviours are larger than what would be expected from the income transfer alone (Macours, Schady, and Vakis 2012). Red Protección Social (RPS), another CCT aimed at improving early childhood development outcomes in Nicaragua, improved male's achievement on cognitive assessments at age 10, but only if they were treated before turning one year old, compared to those who were treated between 1 to 2 years old (Barham, Macours, and Maluccio 2013).

The goals of this paper are to, first, quantify the impacts of Bono 10000 on young children (those children not subject to the CCT education conditions due to their age) and explore the potential channels of impact that may be operating (maternal labour force participation; household expenditures and maternal empowerment and depression); and second, determine if a CCT that incentivises school enrolment in 6-18 year-olds who have not yet completed 9 years of education can spur human capital spillovers from those older siblings to their younger sibling (0-5) through the older sibling attending school longer due to the CCT (Benedetti, Ibarrarán, and McEwan, 2016). The paper's main contribution is this focus on the role policy has in spurring early childhood development through direct or indirect means. Evidence of an intra-household spillover would identify positive early childhood development effects which extend beyond the stated goals of a policy programme. To achieve these goals, we use data from the randomised evaluation of Bono 10000, a CCT in Honduras that distributed cash to poor and extremely poor households. Bono 10000 incentivised households to send older children between the ages of 6 to 18 to school by providing a per household cash transfer. Benedetti et al. (2016) analyse the effects of Bono 10000 on the beneficiary children, finding that the programme resulted in increased school attendance of approximately four percent, while child labour participation slightly decreased. An important result from their analysis was that whether the conditionality was binding or not mattered in the effectiveness of the transfer. As we will explain below, households with two or more children of school age were required that at least one of them attended school in order to receive the full transfer, de facto weakening the conditionality. The authors concluded that the programme had little to no effect on school attendance in larger households, presumably because the family was aware that only sending one child was enough to receive the full benefit. Large households are likely to have older children. Older children have a higher opportunity cost of attending school, so a reasonable response by households was to send the youngest to school (Benedetti et al. 2016). Also, households with more eligible children are poorer, a stylised fact in Latin America.

This paper examines the effects of the CCT on the development of children 0-60 months old and provides a plausible mechanism as of why these impacts take place. We argue that the impacts occur mainly because of sibling interactions in the household. Specifically, we advance that sibling spillover effects may explain how effects flow from an older sibling to the younger sibling, as a result of their relationship within the household.[21] In this situation, older siblings with increased educational attain-

[21] We also considered two additional hypotheses of why we observe these effects: firstly, behavioural change can be because the programme did some labelling of the transfers inducing changes such as larger consumption of some items (books, play material, etc.), then provoking better interactions and therefore better communication skills in children. However, because there was no social marketing

ment should have higher levels of human capital which enable them to communicate better with their younger siblings. This suggests that their relationships with their siblings will then increase in quality, as they are better able to relate to their siblings, boosting the younger sibling's development. The existing literature provides evidence that these spillovers are possible for academic achievement in a non-experimental setting (Nicoletti and Rabe 2014; Nielsen and Joensen 2015; Adermon 2013). The presence of these spillovers is important in relation to the extensive research attempting to determine when investment in human capital has the highest returns (Cunha and Heckman 2007). Early childhood development investment can have multiplier effects later in terms of skills formation, which leads to better outcomes for children as they age. It is important to identify and measure the size of these indirect effects as they can help promote the importance of these programmes as cost-effective and comprehensive tools for poverty reduction.

A few papers in the literature have addressed sibling effects, but only in terms of enrolment of the younger sibling. We are not aware of any study looking at early childhood development (ECD) outcomes that has found significant indirect effects from a CCT. Using this unique RCT, we show that there are improvements in younger sibling's human capital because of Bono 10000, measured by scores on the ASQ test administered to children under the age of 5. The programme improves children's communication skills, especially in households where there are two to three older children, and when the children subject to the conditions have between four and six years of education. The spillover effects are greatest in children between the ages of 25 and 36 months, a key period in the development of a child. The results are robust to false rejection of the null hypothesis through the Romano-Wolf (2005) multiple hypothesis testing framework. The mere presence of indirect effects shows

nor labelling whatsoever, we can rule out this effect. Table C.8 shows some significant (but very small) differences in protein and vegetable purchases, and a larger significant difference in whether the household reported the Bono has been spent in food, health and education. However, Table C.9 shows that the latter variable is not behind the impact on ASQ. Another hypothesis is that due to the additional cash mothers are now less stressed and have more bandwidth to properly interact with their children, Table C.11 also shows that there is no impact on children due to the interaction of treatment with these variables (and the sample is different as not all mothers responded this part of the questionnaire). Moreover, Table C.10 shows that there is no impact on maternal labour status, implying no additional time with children either.

that spillovers must be considered when policy makers are designing their programmes.

The rest of the paper continues in the following way. Section 4.2 presents a literature review on sibling effects. Section 4.3 presents the Bono 10000 CCT in greater detail leading into Section 4.4 which develops the methodology that is used in the analysis. Section 4.5 presents the data with Section 4.6 presenting the full results of the study with concluding remarks.

4.2 Literature Review

Several recent studies have focused on different types of sibling spillovers due to different programmes ranging from providing school scholarships to food sharing mechanisms. The basis of the literature comes from an intersection in the economics and psychology research, where the psychological bond between siblings can possibly lead to improved outcomes for their younger siblings. Summarising the literature on the psychology of sibling dynamics, Dunn (1983) discusses many early studies which show that older children with higher cognitive ability are more effective at helping their younger siblings with less cognitive ability solve problems as they exploit their intimate relationship. Smith (1990) also shows evidence that when older children have a larger vocabulary they are more likely to teach the younger sibling. With this as a starting point, Nicoletti and Rabe (2014), Nielsen and Joensen (2015), and Adermon (2013) establish relationships between siblings' academic attainment and achievement. Nicoletti and Rabe (2014) show that older siblings (of about 14 to 16 years old) have a direct effect on their younger sibling's test scores in UK middle school examinations of a magnitude of 4%. The effect is double the size in poor neighbourhoods. Results suggest that these older siblings may be developing skills in school that make them more effective teachers at home than their parents, who have lower levels of education. Nielsen and Joensen (2015) show that older siblings who were affected by a pilot programme incentivising math and science courses in Denmark were more likely to have younger siblings enrolled in these same courses, even though the incentive to take them had been removed by a later change in the school system's policy. Finally, Adermon (2013) shows that the education attainment of older siblings is an important factor that is positively related to their younger sibling's education attainment.

Other studies have explicitly focused on the heterogeneity in sibling relationships in terms of gender and age differences. In terms of gender, an older sibling may have different responsibilities in the household such as older sisters being responsible for caring for the younger siblings in the household, while older brothers are more responsible for supplementing incomes in the household. Qureshi (2011) focuses on the impact of older sisters in Pakistani households. In this setting, the oldest sister has a large share of the responsibility of care for her younger siblings. The author shows that an older sister with higher education attainment leads to younger siblings that are more likely to be enrolled and literate (up to 19% higher probability). Further, the age difference between siblings can describe the way they interact. Much older siblings may act in a distant role by not adapting their language to their younger siblings (Dunn, 1983; Harkness, 1977)[22] which does not benefit the younger sibling's development. The evidence in Nielsen and Joensen (2015) and Adermon (2013) supports both conclusions with variable effects depending on brother-brother or sister-brother relationships and smaller effects for large differences in age. Lastly, other studies focus on the effects of a disabled sibling, gender and how birth order influences outcomes (Behrman and Taubman 1986; Kessler 1991; Lindert 1977; Fletcher, Hair and Wolfe 2012).

There are a few examples of research that accounts for indirect sibling spillovers in human capital from CCT's, yet do not come to conclusive results on their existence. Ferreira, Filmer and Schady (2009) study a Cambodian CCT that provided scholarships to middle school aged children who were at risk of dropping out of school. Although the direct effect was positive, the indirect effect on siblings in the household was inconclusive. Barrera-Osorio, Bertrand, Linden and Perez-Calle (2008) and Lincove and Parker (2016) focus their studies on Latin American countries, which are more relevant to the Honduras context. In the first study, the authors use a Colom-

[22]Harkness (1977) notes the difference between the way mothers and siblings speak to the younger children in the household, with mothers continuously asking questions of the child to elicit responses from the child, whereas the siblings commentate the world around them, which means that the younger sibling does not speak as much (Dunn [1980], p. 796).

bian CCT to show that the positive direct effect is met with a significant negative indirect effect that stems from the crowding-out effects in households. Untreated children are less likely to be attending school and more likely to work, showing that households may direct resources away from siblings in the household and towards the beneficiary child's human capital development. Lincove and Parker (2016) study the effect of the Red de Protección Social CCT in Nicaragua across children depending on their eligibility status for the transfer. While the effectiveness of the programme on the eligible children varied across their status, the indirect sibling effects were always insignificant. We aim to build on this existing gap in the literature on indirect effects of conditional cash transfers through spillover effects.

4.3 Bono 10000

Bono 10000 was a CCT introduced in Honduras in 2010 as an expansion of the previous CCT in the country, PRAF-II. PRAF-II was the continuation of the Honduran Government's existing poverty reduction efforts which distributed cash to poor household's dependent on them meeting certain school enrolment conditions for older children and using medical services in younger children. Bono 10000 was the next modification by the government to improve these efforts. The value of the transfers was increased in addition to the eligibility constraints being changed. Most importantly, the transfer went from being a per-child to a per-household transfer. To be eligible for the programme, households had to: i) reside in a village declared eligible, and ii) be poor as defined by a proxy means test. Households were to receive the transfer through three instalments paid out by the government each year. There were two different levels of transfers in Bono 10000 based on households meeting conditions that designated their use of education and health services in their local area. The larger transfer, worth 10,000 lempiras (US$500), was conditional on school attendance for children between the ages of 6 to 18. Children were required to enrol in school and attend classes regularly (80% of the time). In households that had two or more children in school age, the programme only required that one of them meet the conditionality in order to receive the transfer. A smaller transfer of 5,000 lempiras (US$250) was associated on infant health and was applied to children ages 0-60 months. In this case, the conditionality was that the child should comply with the mandated number of check-ups based on their age. This conditionality applied only to households that did not have any children of school age, for which the larger educational transfer applied. Also, if there were two children in the 0-60-month age group, if only one of them attended regular health check-ups the household would receive the transfer. This structure of the transfers resulted in that more than 80% of the households and 90% of the resources were linked to the educational side of the programme.[23] Hence, in this paper we focus on how the educational component of the CCT may have generated impacts on the child development of the younger children.

The transfer amount being defined per-household rather than per-child has some key implications. Larger households receive smaller per capita transfers as a result, leading to different behaviours depending on the size of the household. These households needed to enrol one child in school which decreased the likelihood of the cash transfer inducing new enrolees in larger households (because one of their siblings likely was enrolled already) (Benedetti et al. 2016). The transfer was approximately 18% of average per capita income and was larger than the previous CCT's in the country (Galiani and McEwan 2013). Spillover of the cash transfer was minimal, with only 7% of the control group receiving the transfer before the follow up (after which everyone in the control group was supposed to receive the money), and 91% of the treatment group receiving the transfer for approximately 9 months before the follow up survey.

[23] As households could only receive one of the transfers (and were automatically disqualified for the lesser amount if they had an older child), an interesting discussion is which of the two transfers had higher effects on the 0-5 children (even of the beneficiary household will have a different composition) We had only 383 children in the sample who meet the criteria (under age 5 and no older children) for the smaller health transfer of $250. A regression was run on this subsample, finding no effects on ASQ from this smaller transfer.

4.4 Methodology

4.4.1 Sample Selection and randomisation

Initially, the government of Honduras began implementing the transfer before the randomisation occurred, with a focus on the poorest aldeas (communities) in Honduras. As a result, the surveyors were unable to use them and had to select 816 less poor aldeas as their experimental group, randomising amongst this group to choose 150 treatment and control aldeas each. The process is described in Figure 1. Figure 2 shows the location of each of the aldeas throughout Honduras. In each of the 300 selected aldeas, surveyors from NORC selected 15 households to collect data from in a survey distributed in early 2012. A weakness of the survey procedure is the fact that the transfers began while the baseline survey was still occurring. This is discussed further in Benedetti et al. (2016) who found that households who received the money early were not significantly different from those who received the transfer on time. Once this process was completed, there were a total 4416 households and 296 rural aldeas interviewed in January-June 2012 and March-June 2013[24]. The surveys included a range of questions focusing on the assets and quality of the household, characteristics of individual members ranging from their education to their income, and health for mothers and their children.

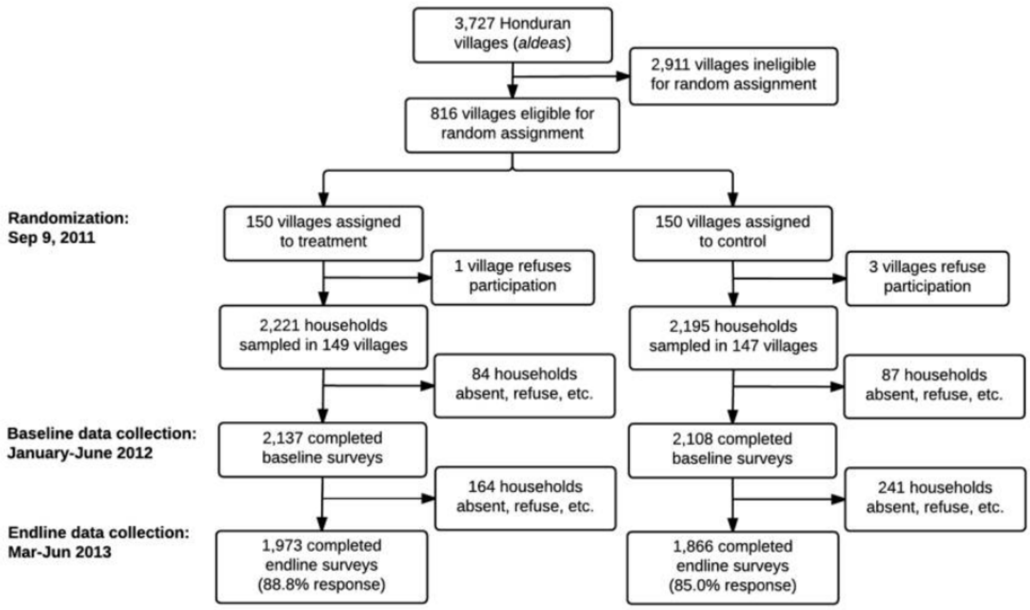
Attrition bias is a slight concern in this situation. Benedetti et al. (2016) argues that for the entire sample of children from 0 to 17-years-old it should not be a significant concern.[25] Table A.1 performs a similar analysis, presenting the differences in means for several different household and individual characteristics by treatment group for the different subsamples of households in the survey. The results are similar to those in Benedetti et al. (2016), except in our sample there are significant differences in households having piped water, access to telephone services, household size, and access

[24]Four of the villages, one in the treatment group and three in the control group, declined to provide any answers for the survey. Attrition is minor in the survey, with 88.8% of households in the treatment group providing a response in the follow up, and 85% in the control group. The percentages drop slightly to 88.2% and 83.3% when the non-responding villages from the initial randomisation are considered. This is seen in Figure 1 (Benedetti et al. [2016], p.11).

[25]See Benedetti et al. (2016) for further details on the randomisation method and on the details of attrition.

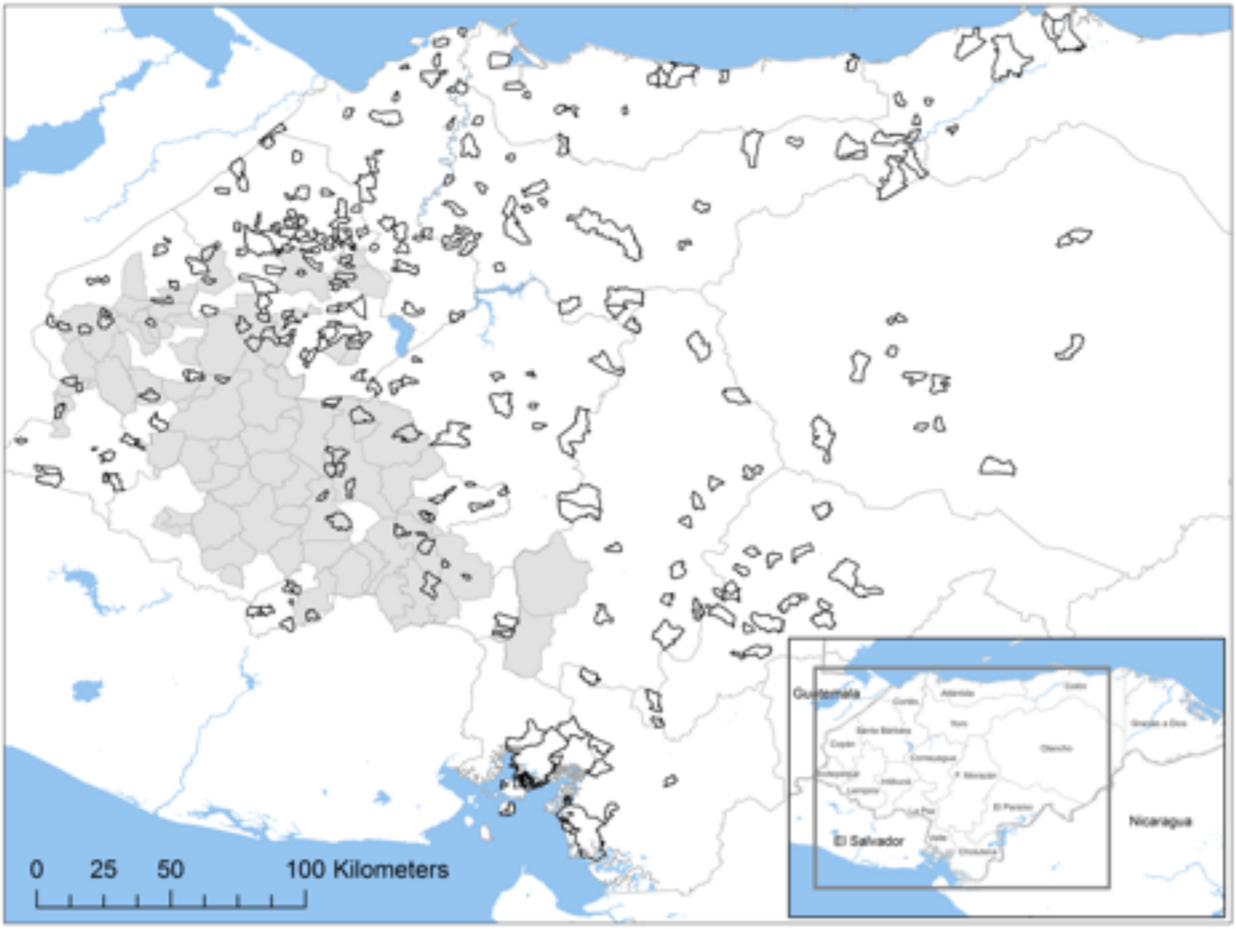
to electricity. While the differences are small, it seems that once the dataset is pared down to having only those households with ASQ measurements, control households are slightly wealthier and smaller. In our estimation, these differences should not play a role in biasing the results. At the very least, we concede that the provided estimates are lower bound estimates. Additionally, upper and lower bounds for effect sizes after a trimming procedure will be presented (Lee 2009) to guard against attrition bias, as is the case in Benedetti et al. (2016).

Figure 4.1: Bono 10000 Randomisation Procedure



Source: Benedetti et al. (2016)

Figure 4.2: Sample Territories in Bono 10000 and PRAF-II



Notes: Shaded areas indicate 70 municipalities sampled in the PRAF-II experiment in 2000 (Galiani & McEwan 2013). Outlined areas indicate 300 villages (aldeas) sampled in the Bono 10,000 experiment in 2011. Source: (Benedetti et al. 2016)

4.4.2 Child Development measure: Ages and Stages Questionnaire (ASQ-3)

The Ages and Stages Questionnaire (ASQ-3) is a screening test for measuring children's development. The test can be used for children between the ages of 1-66 months, with different questionnaires of the test designed for different age brackets of the child. The test is available in Spanish from the publisher (Squires and Bricker 1999), is self-administered by the mother, and takes no more than 10 to 20 minutes to complete. There are 30 questions across five different domains covering communication (vocalisation, understanding), problem resolution (measured by playing with toys), social skills (playing by themselves and with others) and executive function (measured by the ability to move limbs).[26]

The ASQ test is self-reported by mothers recalling their child's behaviours in the preceding time before the test. It is a reliable and low cost measure of early childhood development, previously being shown to have a positive correlation in Chile (Schonhaut et al. 2009) with another test of early childhood development, the Bayley Scale of Infant Development (Bayley-III). One of the attractive characteristics of the ASQ is that it provides a comprehensive picture of early childhood development. While measurement error is a possible concern in the test scoring, the test has been used and proved useful in several different contexts (Fernald, Engle, and Raikes 2009; Martinez, Naudeau, and Pereira 2012) and the ease and low cost of administration make it an attractive tool to measure early childhood development.

The main ASQ measure and the individual domains can be used for analysis in three different ways; a raw score or normalising either on a reference scale or within the sample. The raw score is simply the number of points a child receives on the test. There are three possible responses for each item on the test. First, if children always exhibit a behaviour or movement described on the questionnaire, they receive a mark of 10. 5 points are given if the child sometimes exhibits or performs one of the described behaviours. Finally, 0 points are given if the child never behaves or makes movements in that way. The maximum total score then is 60 for each domain, and 300 overall. Since the fine motor skills are not measured in this survey, the maximum score is 240.

Next, there are the two methods of standardising the test scores based on the age of the child. The within standardisation normalises the test scores based on the means and standard deviations of the sample, which is useful for determining a local estimate of how the cash transfer is impacting children's outcomes. The external normalisation

[26]The fine motor skills measurements were left out of survey due to the higher measurement costs from increasing the length of the test. By leaving the measure out, it is possible to look at problem resolution skills, which can act as a robustness check on communication (because they are cognitive in nature). It is important to note that this measure for children who are 15-20 months old is downward skewed because there are only five questions for problem resolution, rather than six. This affects the ability to compare this sphere across ages because it is impossible to get the same total score.

is based on an international standard and allows the children to be compared to other children around the world, rather than just their counterparts on the study (Fernald et al., 2009). The standardisation is performed using norms that have been previously set out (Squires and Bricker, 1999). To test the reliability of these measures, we use Cronbach's alpha, a statistic that is used in psychometric literature[27] for the raw scores. For each domain, the measure generally is higher than 0.7, except for total ASQ, which is found to be 0.67. The alpha for the total measure is 0.78, meaning that we can be confident of the variables internal validity in measuring the domain of interest.

We expect that the cognitive domains such as communication and problem resolution are the ones that should be affected most by sibling spillovers. One of the major benefits of schooling is that children interact with peers in close quarters, dealing with problems and communicating their feelings to each other. By increasing schooling attainment, children not only gain additional skills in numeracy and literacy, but also in dealing with others. They then take these newfound skills to the household, where they are better able to communicate, empathise and teach their younger siblings, fostering this spillover of skill. Motor skills are less impacted by the additional level of schooling in the household, as these skills are developed early on and fine-tuned later in life.

It is important to note that this is the first time that ASQ-3 has been administered in Honduras. Tables C2a-C2d in the appendix show baseline scores by age and by domain. Results in raw scores are well aligned with international studies (ENDIS 2015; Berlinski and Schady 2015) with most raw scores being near 50 points. There is one less question for 12-23 month olds in the problem resolutions domain which explains the significantly lower scores across all the subsamples. The next two columns show standardised scores, discussed further in the next section. We first show scores based on the internal standardisation, and then based on external standardisation.[28] Comparing the sampled children with the international standard (the last two columns)

[27] Acceptable measures for reliability range from 0.7 and up (Cronbach 1951)

[28] External standardisation is set on an international scale. The process is made possible for Spanish-speaking countries by the previous work of Squires and Janson (2004) and Squires et al. (2008).

highlights how the large gaps in socioeconomic status can negatively affect the human capital development of children early in their lives. The tables show that children's communication, motor, and socio-individual skills lag behind their international counterpart's skills by approximately 0.25 standard deviations across the different subsamples. In terms of problem resolution, the children are disadvantaged even further, with the difference being nearly a full standard deviation. These tables paint a picture of the overall lack of development which necessitates some sort of intervention in the short-term to improve outcomes.

4.4.3 Subsamples

The baseline data is comprised of 2935 children younger than 5 years old in the 4416 households. 2520 of these children took the ASQ test at the baseline. For the analysis, we break this number down into sub samples to aid in identifying the sibling spillover effects.

a) Full sample: 2935 children in 4416 households

b) Full panel: 1505 children with baseline and follow up ASQ data (Sample 1)

c) Full panel with siblings of school age: 1122 children (Sample 2)

d) Full panel with one or two siblings of school age: 715 children (Sample 3)

In the largest subsample, we include only those children who have fully completed the ASQ test at baseline and follow up, leaving 1505 children. This sample is used in the initial regressions to provide estimates of the overall effects of the programme. To study the channels of these effects and particularly heterogeneous effects amongst households with or without an older sibling (and therefore beneficiaries of the larger transfer), we purge the sample of households where there are no school aged children. This is Sample 2, where there are 1122 children. Sample 3 provides an additional layer of heterogeneity, and removes those households with more than two older children (following Benedetti et al. [2016]), leaving 715 children in the panel. This latter subsample represents a direct link between the beneficiary child and their younger sibling, which is the key in determining whether there are within-family sibling spillovers. Sample bias is a concern here, because these households are compositionally different than the other subsamples. Households with one to two eligible children are generally younger and smaller. The older siblings are more likely to be in earlier grades of school, which could mean that their impact on the younger sibling's human capital is diminished because they both have similar stocks of capital. This further supports the conclusion that our estimates are lower bound estimates of the true spillover effect.

4.4.4 Identification Strategy

In general, endogeneity of the estimates of an intervention on child development might be a concern because of unobservable family and child characteristics. These range from children's endowments to the characteristics of the household such as how many other siblings there are and what roles do the parents believe that their children should play depending on their gender. However, the successful randomising of the evaluation sample in this programme takes care of these concerns because we can consider the treatment as orthogonal to unobservable household characteristics. It also means that since the structure of the groups is the same, we can look directly at the interaction between the treatment and a few different heterogeneous effects to estimate the programme's impact on the human capital development of younger siblings due to their older siblings increased school attainment. Following Benedetti et al. (2016), we use the following empirical specification:

$$Y_{ij} = \alpha_0 + \alpha_1 T_v + \alpha_2 \mathbf{X}_{ij} + \alpha_3 \mathbf{E}_i + \epsilon_{iv} \tag{4.1}$$

In this equation, is the z-score of the total ASQ (or one of the individual domains), is the treatment status of the child (whether they are in a treated aldea or not). These include eligibility status of the household, number of enrolled children, years of education for the older sibling, gender of the older sibling, and age difference. X_{ij} is a matrix that includes an extensive number of different household and individual controls which are listed in Table 4.1. In the cases where the heterogeneity is categorical (such as grades), the regression takes the following form:

$$Y_{ij} = \alpha_0 + \sum_{j=1}^{n} \alpha_{ihj} T_v \mathbf{E}_{ihj} + \alpha_{n+1} \mathbf{X}_{ij} + \alpha_{n+2} \mathbf{E}_i + \epsilon_{iv} \tag{4.2}$$

The summation term allows for flexibility when analysing different levels of the interaction of interest. J denotes the number of levels of whatever variable is being used for $\mathbf{E}_{ih}$. For example, since grade enrolment is broken into terciles, j would be three. In both equations, we use robust standard errors, clustered at the aldea level.[29] The estimate of α_1 (or α_{ihj} in Equation (4.2) for child i, household h, and heterogeneous interaction j) is an intent to treat (ITT) OLS estimate (Angelucci and Di Maro, 2016), because the treatment variable is based on eligibility rather than actual receipt of the transfer. The estimate can be interpreted as the change in the younger sibling's human capital development due to the marginal change in the older sibling's value of the studied characteristic.

The regressions use the three subsamples described previously. The preferred specification is when the sample is limited to more than one eligible child between the age of 6 and 18. This is because the transfer could cover the level of income that is brought in by the oldest child in the household, thus making it more beneficial for households to send the youngest child to school. However, this is not always the case, especially since the amount of the transfer stays constant no matter how many children are enrolled in school.

4.4.5 Romano-Wolf Multiple Hypothesis Testing

A potential methodological issue that arises when testing the statistical significance of the treatment variable across the multiple domains of the ASQ is that there is a

[29] There are 291 aldeas, making this a very useful process.

possibility of Type-1 error from chance. Some hypothesis tests will appear significant when in fact they are not. To combat this issue, we make use of the Romano-Wolf (2005) method, which performs a stepwise test of hypotheses, controlling for the family-wise error rate using specifications of a CAPM model in finance. Carniero and Ginja (2014) advance this method into multiple testing in human capital development measures. In this case, the family-wise error rate can be defined as the probability of incorrectly accepting that one specification of our model is better than the others. By controlling for this in our hypothesis testing, we adjust the p-values to the family-wise error, making it harder to achieve conventional statistical significance, and as such, we present a more robust result.

A key feature of the method is that it uses a studentised t-distribution for the critical values. The authors support the selection of the distribution with three arguments. First, they state that the studentised distribution, compared to basic methods, have reasonable power characteristics. Second, the bootstrapping method provides some asymptotic benefits. Lastly, the size of the coverage probabilities is comparable to the basic method and allows for comparison across different parameter estimates since they occur under the same distribution (Romano and Wolf 2005, p. 1254-1255).

Algorithms 4.1 and 4.2 in Romano and Wolf (2005) introduce the test statistics, how to obtain the critical values and outlines the stepwise procedure. The test statistic which we obtain is $z_{R,n} = \frac{\hat{\beta}_{R,n}}{\sigma_{\hat{R},n}}$, where R is the sample size, and n={1,. . . ,N} is the number of hypotheses to test.[30] This estimated parameter,$\hat{\beta}_{R,n}$, is the effect of Bono 10000 on a younger sibling's score in one of the individual domains of the ASQ test.. Next, we create a bootstrapped data matrix, X_T^*, and with the standard deviation of our initial estimator and this bootstrapped data, we estimate a new coefficient of the programme impact, $\hat{\beta}^*_{R,n}$ and obtain a new test statistic $z^*_{T,n} = \frac{\hat{\beta^*}_{R,n}}{\sigma^*_{\hat{R},n}}$. The critical value for this new test statistic, $\hat{d}_j$, is data dependent and obtained through use Algorithms 4.1 and 4.2.

The stepwise testing procedure works as follows. Test statistics are obtained and are labelled in descending order. P-values are obtained from each of the hypothesis tests,

[30] In this case four, for each of the domains of the ASQ test (we do not test the ASQ results).

rejecting each hypothesis of $\hat{\beta}_{R,n} = 0$ if $|n| > \hat{d}_1$, where $\hat{d}_1$ is the critical value from the entire sample. If none of the hypotheses are rejected, the method stops. If at least one test is rejected, that specific test is removed from the process, and the remaining data is used to obtain $\hat{d}_2$. The procedure moves to the next iteration at a given significance level until no tests are rejected.

In our application, the method is used to test the significance on the estimates of the treatment effect and the treatment effect interacted with the dummy for whether the household was eligible to receive the transfer. The procedure is completed over 1000 repetitions across the four domains of the ASQ test. We use only the preferred specifications to capture programme effects with the full set of controls. Testing in the specifications with the additional heterogeneity of categorical factors becomes data intensive, and is not completed.

4.5 Data

Table 4.1 presents the descriptive statistics across the entire sample and the four different subsamples based on the number of eligible children in the household. At baseline, sampled children are on average two and a half years old (but are two on average in the panel sample). Total income is about 5,000 lempiras and household size is, on average, around six. Smaller households are richer than the rest of the sample, following the stylised fact. Mothers have approximately 4.5 years of education, which is nearly half a year more than children's fathers. They also have a higher literacy rate compared to the father's (approximately 84% to 74%) across the different subsamples. In terms of assets, the descriptive statistics provide a picture of the level of poverty in the sample: less than 20% of households have access to piped water, 60% of households have electricity, and 75% have a connected bathroom.

Table 4.2 shows clear gradients by income quintile, which are the most pronounced in the Communication domain (0.17SD), while Figure 4.3 shows very clear gradients by maternal education in the ASQ test and all of the domains for ages 5 and younger.

Table 4.1: Baseline Household characteristics by Subsample and Panel

	(1) All children with ASQ scores at baseline		(2) At least 1 Older Sibling		(3) At Least 1 Eligible Older Sibling		(4) 1-2 Eligible Older Siblings		(5) Only 1 Eligible Older Sibling	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Treatment and Control	0.50	0.50	0.52	0.50	0.52	0.50	0.53	0.50	0.53	0.50
Age in Months	29.01	17.00	22.94	13.46	23.37	13.50	23.93	13.52	24.23	13.74
Total Income	4848.98	9782.88	5013.22	11904.00	5487.07	13568.17	5679.38	14654.63	5652.37	15313.25
Household size	5.96	2.69	5.80	2.55	6.50	2.52	6.71	2.50	6.19	2.26
Female	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50
Mother's years of education	4.62	3.20	4.66	3.13	4.40	3.06	4.38	3.04	4.47	3.09
Mother is literate	0.85	0.36	0.86	0.35	0.84	0.37	0.84	0.37	0.85	0.36
Father's education (years)	3.87	2.79	3.97	2.83	3.68	2.71	3.71	2.67	3.77	2.70
Father is literate	0.77	0.42	0.77	0.42	0.76	0.43	0.77	0.42	0.76	0.43
# of children age 0-5	1.69	0.83	1.67	0.80	1.71	0.84	1.71	0.85	1.67	0.85
# of children aged 6 to 18	1.84	1.62	1.75	1.56	2.35	1.36	2.55	1.39	2.06	1.06
Dirt floor in dwelling (1/0)	0.38	0.48	0.38	0.49	0.38	0.49	0.37	0.48	0.35	0.48
Piped water in dwelling (1/0)	0.16	0.37	0.17	0.37	0.16	0.37	0.17	0.37	0.18	0.38
Dwelling has bathroom or latrine (1/0)	0.73	0.44	0.73	0.44	0.74	0.44	0.75	0.43	0.76	0.43
Dwelling has electricity (1/0)	0.63	0.48	0.64	0.48	0.63	0.48	0.63	0.48	0.66	0.47
Access to telephone (1/0)	0.82	0.39	0.83	0.38	0.83	0.37	0.83	0.38	0.83	0.38
Rooms in Dwelling	3.18	1.43	3.16	1.43	3.19	1.41	3.22	1.41	3.21	1.40
Adults in household who are Lenca (1/0)	0.05	0.21	0.05	0.20	0.05	0.21	0.05	0.22	0.05	0.21
Dwelling only accessible by footpath (1/0)	0.32	0.47	0.30	0.46	0.31	0.46	0.32	0.47	0.29	0.46
Observations	2520		1505		1122		929		715	

Note: Column 1 includes all children in the sample.

Column 2 includes all children who have an older sibling between the age of 6 to 18 and panel data, regardless of eligibility.

Column 3 includes households who have at least one child between 6-18 years old who is enrolled lower than the 9th grade and in the panel.

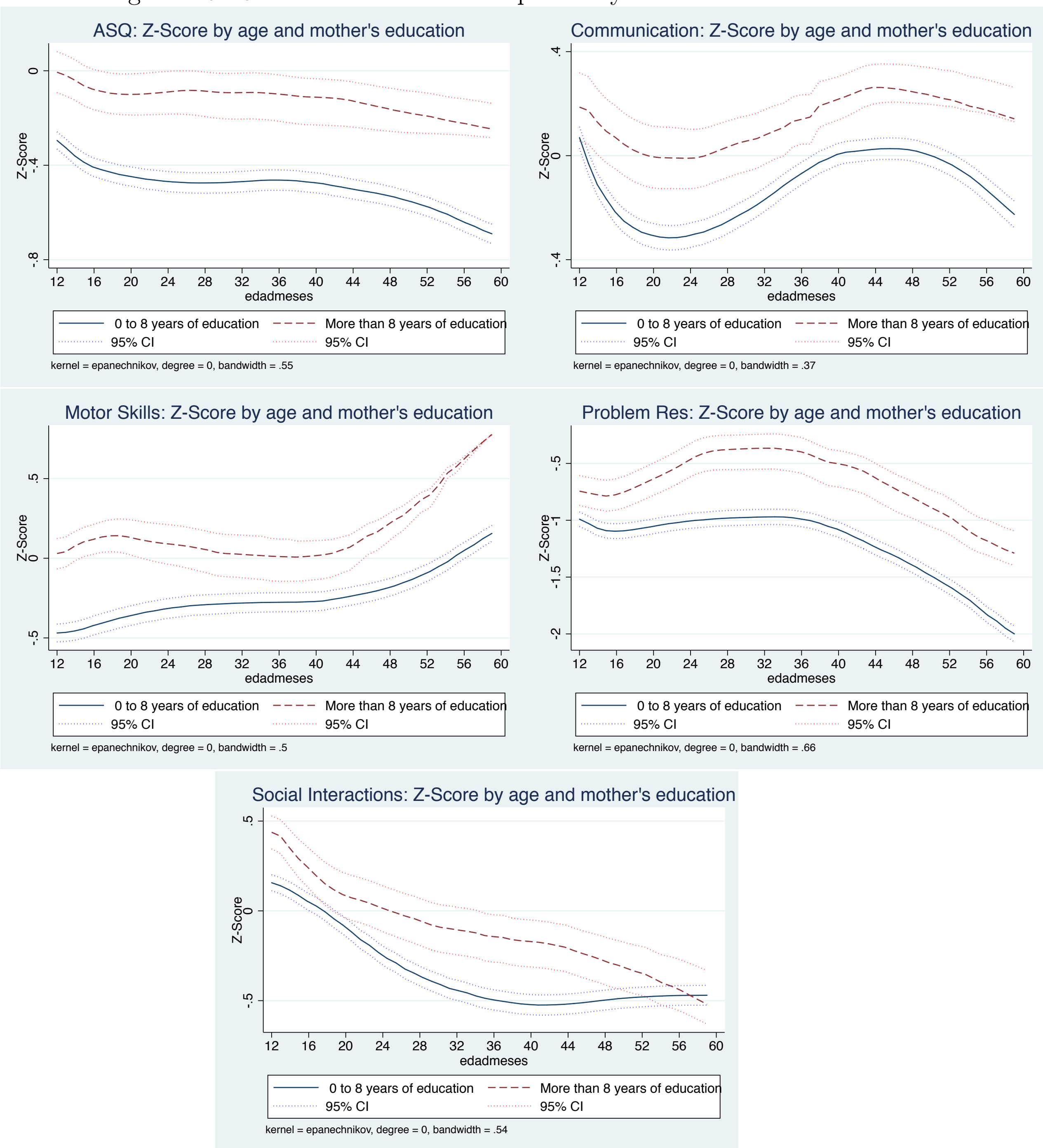
Column 4 includes households with one to two children who meet the constraint and are eligible and in the panel.

Column 5 includes the households where there is only one eligible older child and in the panel.

Table 4.2: Baseline ASQ by income quintile (SD)

	Q1		Q5		Diff (Q5-Q1)
	Mean	SD	Mean	SD	
Communication	-0.07	1.03	0.10	0.87	0.17**
Gross motor skills	-0.02	1.00	0.12	0.90	0.13*
Problem Resolution	-0.04	1.05	0.11	0.97	0.15*
Social Skills	-0.06	1.05	0.10	0.95	0.16*
ASQ	-0.07	1.05	0.14	0.89	0.20**

Figure 4.3: Gradients in child development by maternal education



4.6 Impact Estimates

Tables 4.3-4.8 present the central results of our paper. They show ITT estimates of Equations 4.1 and 4.2 of the effect on the child development outcome of eligibility for Bono 10000. These tables include an indication of whether the null of no effect is rejected when we account for multiple hypothesis testing. Throughout our discussion, we consider that the programme has a statistically significant effect on child development only in the cases where we can reject the null that the effect is zero using the procedure introduced by Romano and Wolf (2005).

We show only the specifications in which we include all the child and household's controls as listed in Table 4.1,[31] as well as the baseline ASQ (or individual domain) score and a variable for time effect. The time variable is based on when the household completed its interview (1 if after March 2013, 0 if before). Moreover, we show only the impacts on the overall ASQ score and the communication domain. The impact on the ASQ is important as it measures the overall construct of child development, while the communication domain is the most likely domain to be affected by the presence and interaction of an older sibling (our main hypotheses). It is also the only one that survives our Romano-Wolf hypotheses testing; and, therefore, for the sake of saving space, the non-significant impacts on the other three domains of the ASQ are presented in the Appendix.

Table 4.3 presents four panels.[32] Panel A presents the results of Equation 4.1 for different subsamples: there is a significant positive effect of the programme only on the communication domain of 0.13-0.14 SD which increases to 0.17SD in the sample of households with 1-2 older siblings. In panel B and all the following regressions, we run Equation 4.2. There we ratify the importance of being eligible for Bono in the first column; as the interaction of treatment and actual eligibility for Bono shows a 0.17SD positive effect on communication domain. Notably, while it is a larger point estimate, it is not significantly to the estimates on the other samples.[33] Finally, Panel C shows

[31]In the regressions, all specifications control for age of the older sibling and the number of siblings between 6 to 18 years old to consider the concern that larger households will not be affected by the constraints since it is a per household transfer.

[32]Table C.4 shows the same specification of Table 4.3 but includes all the 4 domains of the ASQ.

[33]Results available upon request

no heterogeneity of impacts by income quintile, and Panel D shows the impact of years of education of the older eligible sibling. An additional year of education of the older child has 4% of 1 SD effect on their younger sibling's ASQ. Table C.3 in the appendix presents Lee (2009) trimmed upper and lower bounds of the mean level in treatment and control groups. The bounds for communications show positive, non-zero bounds, meaning that the interpretation of the results is not affected.

Table 4.3: Bono 10000 impact on child development (SD), by subsample

	ASQ	Communication	ASQ	Communication	ASQ	Communication
Panel A: Mean Impact						
Treatment	0.068	0.134**	0.039	0.143**	0.042	0.170**
	[-0.05,0.2]	[0.01,0.3]	[-0.09,0.2]	[0.002,0.3]	[-0.1,0.2]	[0.02,0.3]
RW Hypothesis Rejection at 10%		Yes		Yes		Yes
Panel B: Bono Educación Beneficiaries						
Treatment * Eligibility	0.036	0.165**	0.027	0.159**	0.042	0.170**
	[-0.1,0.2]	[0.02,0.3]	[-0.1,0.2]	[0.009,0.3]	[-0.1,0.2]	[0.02,0.3]
RW Hypothesis Rejection at 10%		Yes		Yes		Yes
Panel C: Treatment by Income Quantile (Base Level is Q5)						
Treatment * poorest quintile	0.105	0.164	0.142	0.131	0.225	0.113
	[-0.1,0.3]	[-0.05,0.4]	[-0.1,0.4]	[-0.1,0.4]	[-0.1,0.6]	[-0.2,0.4]
Q1	-0.127	-0.099	-0.050	-0.028	-0.071	-0.028
	[-0.3,0.03]	[-0.3,0.05]	[-0.2,0.1]	[-0.2,0.2]	[-0.3,0.2]	[-0.3,0.2]
Panel D: Treatment interacted with years of education of older sibling at baseline						
Treatment * Years of Education	0.017	0.032**	0.016	0.032**	0.024	0.044**
	[-0.01,0.04]	[0.002,0.06]	[-0.01,0.04]	[0.002,0.06]	[-0.01,0.06]	[0.01,0.08]
Household Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time and Lagged ECD	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1505	1505	1122	1122	715	715

Note: Standard errors clustered by Aldea. The first sample includes all households in the panel.
Sample 2 is all the households in the panel with at least one older sibling. Sample 3 is the households with at least one eligible older sibling,
All regressions include a constant. Model includes a selection of household and individual controls found in Table 4.1
as well as time, baseline ECD measures, and department fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.4 shows a falsification test to test whether the effects are because of the programme, and not wider trends in the country or early childhood. We test this by examining four specific subsamples; first, a subsample of households that do not meet the larger transfer Bono conditions (i.e. no older children), second, a subsample of households with older children who do not meet the conditions, third, three or more children and fourth, 1 to 2 older children ineligible for the Bono. None of the estimates are significant, showing that the results in Table 4.3 are robust and stem from the programme.

Table 4.4: Falsification Test

	ASQ				Communication			
Treatment	0.090	0.120	-0.009	0.013	0.070	0.082	0.123	0.030
	[-0.1,0.3]	[-0.05,0.3]	[-0.2,0.2]	[-0.2,0.2]	[-0.1,0.3]	[-0.08,0.2]	[-0.1,0.4]	[-0.2,0.2]
Observations	383	576	431	425	383	576	431	425

Note: the first column in ASQ and Communication is composed of those households with no children between 6 to 18,
while the second column are those households with older siblings who do not meet the conditions of Bono 10000.
The third column is households with three or more older siblings and the fourth column are those with 1 to 2 older, but ineligible children
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4.5 (and its analogue in C.5 with all the ASQ domains) expands into the specific household compositions beyond the subsamples. Whereas breaking down the sample into subsamples leads to no significant differences amongst the estimates, the full analysis indicates that there is a sweet spot in households with two children between the ages of 6 and 18. The effect grows once the treatment effect is interacted with two eligible children in the household, identifying a possible spillover effect in younger households.

Table 4.6 digs deeper to find if there is a certain age where these effects are greatest. Looking specifically at the effect of being in a treated area, there is a nearly a full 0.10 SD larger coefficient when children are 25 to 36 months old (from ~0.13 SD to 0.20 SD in the full sample, from 0.14 SD to 0.23 SD in the second). The effect is consistent across the subsamples, even if in the in the smallest sample, statistical power may be behind a less significant coefficient. In the third subsample, the first year of life coefficient is slightly significant and increases the effect up to 0.26 SD in the 0 to 12-month age range. It is possible that the effect is greatest in this period (0-36 months, overall) because it is a sensitive period of development or because of a multiplier effect of children attending crèche and other preschool care programmes.

In Table 4.7, we look at the heterogeneity of impacts, focusing on the two subsamples where there is at least one child older than 6 in the household.[34] This set of subsamples is motivated by both the results seen in Table 4.3 (increasing effects when subsamples

[34]Table C.5 shows the same specification of Table 4.4 but includes all the 4 domains of the ASQ.

Table 4.5: Heterogeneity in Treatment by Household Size

	ASQ		Communication	
Panel A: Treatment status for number of children in the household				
Treatment * 1 child	0.104	0.062	0.024	0.032
	[-0.2,0.4]	[-0.2,0.3]	[-0.2,0.3]	[-0.2,0.3]
Treatment * 2 children	0.160	0.154	0.153	0.177*
	[-0.04,0.4]	[-0.04,0.3]	[-0.06,0.4]	[-0.02,0.4]
Treatment * 3 children	0.097	0.065	0.074	0.047
	[-0.2,0.4]	[-0.2,0.3]	[-0.2,0.3]	[-0.2,0.3]
Treatment * 4 children	0.075	0.016	0.223	0.190
	[-0.3,0.4]	[-0.3,0.3]	[-0.07,0.5]	[-0.09,0.5]
Panel B: Treatment interacted with the number of eligible siblings in the household				
Treatment x 1 eligible child	0.020	0.013	0.113	0.120
	[-0.2,0.2]	[-0.2,0.2]	[-0.1,0.3]	[-0.08,0.3]
Treatment x 2 eligible children	0.185	0.162	0.362***	0.360***
	[-0.06,0.4]	[-0.06,0.4]	[0.1,0.6]	[0.1,0.6]
Treatment x 3 eligible children	0.175	-0.002	0.250	0.143
	[-0.2,0.5]	[-0.3,0.3]	[-0.1,0.6]	[-0.2,0.5]
Treatment x 4 eligible children	-0.046	-0.154	-0.240	-0.248
	[-0.5,0.4]	[-0.5,0.2]	[-0.7,0.2]	[-0.6,0.1]
Household Controls	No	Yes	No	Yes
Time and Lagged ECD	No	Yes	No	Yes
Observations	1505	1505	1505	1505

Note: Standard errors clustered by Aldea. The first sample includes all households in the panel.
All regressions include a constant. Model includes a selection of household and individual controls found in Table 1 as well as time, baseline ECD measures, and department fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

include smaller households) and by the fact that the largest and most significant enrolment effects of the cash transfer were found in households with 1-2 eligible older siblings (Benedetti et al. 2016). This sample restriction allows us to create a link between the one beneficiary child and their younger sibling, introducing the possibility of measuring intra-household spillovers.[35]

Panel A shows that the impact for years of education are larger in the second tercile of the distribution of education of older siblings (4-6 years of education) and stands at 0.28 SD for the communication domain. In the third tercile the effects are also

[35] We cannot observe who the specific beneficiary is (or even if there is one) from the survey, because the survey asks simply whether a child is enrolled or not, not if this child would be directed to labour activities if the transfer was absent. We therefore assume that the oldest child is the beneficiary child in absence of knowing the specific child

Table 4.6: Age Gradients

	ASQ	Communication	ASQ	Communication	ASQ	Communication
Panel A: Age Gradients in Treatment						
0-12 months old	0.083	0.137	0.033	0.112	0.059	0.261*
	[-0.1,0.3]	[-0.07,0.3]	[-0.2,0.2]	[-0.1,0.4]	[-0.2,0.3]	[-0.05,0.6]
13-24 months old	0.024	0.090	-0.007	0.061	0.006	-0.020
	[-0.2,0.2]	[-0.1,0.3]	[-0.2,0.2]	[-0.2,0.3]	[-0.3,0.3]	[-0.4,0.3]
25-36 months old	0.025	0.199*	-0.001	0.232**	-0.016	0.206
	[-0.2,0.2]	[-0.01,0.4]	[-0.2,0.2]	[0.01,0.5]	[-0.3,0.2]	[-0.04,0.5]
37-60 months old	0.149	0.089	0.150	0.145	0.121	0.202
	[-0.09,0.4]	[-0.2,0.3]	[-0.1,0.4]	[-0.1,0.4]	[-0.2,0.4]	[-0.1,0.5]
Panel B: Age Gradients in Treatment x Eligible						
Treatment *	0.034	0.218*	0.032	0.217*	0.059	0.261*
Eligible x 0-12 months old	[-0.2,0.2]	[-0.01,0.4]	[-0.2,0.3]	[-0.03,0.5]	[-0.2,0.3]	[-0.05,0.6]
Treatment *	0.004	0.022	-0.037	0.009	0.006	-0.020
Eligible * 13-24 months old	[-0.2,0.2]	[-0.2,0.3]	[-0.3,0.2]	[-0.3,0.3]	[-0.3,0.3]	[-0.4,0.3]
Treatment *	0.044	0.269***	0.051	0.288***	-0.016	0.206
Eligible * 25-36 months old	[-0.2,0.2]	[0.07,0.5]	[-0.2,0.3]	[0.07,0.5]	[-0.3,0.2]	[-0.04,0.5]
Treatment *	0.047	0.097	0.051	0.068	0.121	0.202
Eligible * 37-60 months old	[-0.2,0.3]	[-0.2,0.3]	[-0.2,0.3]	[-0.2,0.4]	[-0.2,0.4]	[-0.1,0.5]
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects and lag ECD	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1505	1505	1122	1122	715	715

Note: Standard errors clustered by Aldea. All regressions include a constant.
Sample 1 is all the households in the panel with at least one older sibling.
Sample 2 is the households with at least one eligible older sibling, and sample 3 is the households with 1-2 eligible siblings.
Model includes a selection of household and individual controls found in Table 1,
as well as time, baseline ECD measures, and department fixed effects.
Results of the fine motor skills, problem resolutions and social skills are available on request for brevity
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

significant, and are much larger in households with 1-2 eligible siblings, but stay similar in magnitude to the latter category. The education constraint starts to become more binding as children get closer to 6 years of education, possibly explaining these effects. Panel B shows estimates of impacts of 0.20 SD when the eligible child is female, supporting the previous evidence which stated that the relationship between older females in the household could be more mothering. More human capital will increase the quality of the interactions between older and younger sibling, leading to better communication and better early childhood development. In Panel C, we see that one more year of an age gaps between siblings lead to a further benefit of 0.02 SD in young children's communication, again supporting a possible link between a more parental role in older children (Adermon 2013; Nielsen and Joensen 2015; Qureshi 2011). While it could be argued that using baseline education means that this is not

Table 4.7: Heterogeneous Bono 10000 impacts on Child Development(SD), by subsample

	ASQ	Communication	ASQ	Communication
Panel A: Treatment interacted with Terciles of Older Sibling's Education at Baseline				
Treatment *	-0.149	-0.138	-0.207	-0.129
0-3 Years of Education	[-0.4,0.1]	[-0.4,0.1]	[-0.5,0.1]	[-0.4,0.2]
Treatment *	0.033	0.279**	0.054	0.261**
4-6 Years of Education	[-0.2,0.2]	[0.05,0.5]	[-0.2,0.3]	[0.04,0.5]
Treatment *	0.141	0.174*	0.209**	0.267**
7-8 years of education	[-0.04,0.3]	[-0.02,0.4]	[0.0003,0.4]	[0.03,0.5]
Panel B: Treatment interacted with Gender of older sibling (Female equal to 1)				
Treatment *	0.075	0.201**	0.057	0.202*
Gender	[-0.09,0.2]	[0.01,0.4]	[-0.2,0.3]	[-0.03,0.4]
Panel C: Treatment interacted with Age Difference of Older Sibling				
Treatment *	0.006	0.013*	0.008	0.020**
Age Difference	[-0.01,0.02]	[-0.001,0.03]	[-0.007,0.02]	[0.003,0.04]
HH Controls	Yes	Yes	Yes	Yes
Time Effects and lag ECD	Yes	Yes	Yes	Yes
Observations	1122	1122	715	715

95% confidence intervals in brackets. Note: Standard errors clustered by Aldea. All regressions include a constant.
The first sample includes all households in the panel. Sample 2 is all the households in the panel with at least one older sibling.
Sample 3 is the households with at least one eligible older sibling, and sample 4 is the households with 1-2 eligible siblings.
Model includes a selection of household and individual controls found in Table 4.1 as well as time
, baseline ECD measures, and department fixed effects.
Results of the fine motor skills, problem resolutions and social skills are available on request for brevity
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

evidence of a spillover, we argue that it is a simple step to say that more education is going to improve relationships and increase the programme effect.

Table 4.8 builds on the gender effects found above by focusing on brother or sister sibling pairs (i.e. if the younger sibling is male: seeing the effects on whether the older sibling is male or female, and same if younger sibling is female). There does seem to be an important consequence of having older females in the household when the younger sibling is a boy, with larger effects of 0.25 SD in households with at least 1 eligible sibling. Again, this is likely due to the parenting role of older sisters, particularly with boys that are those that show the largest deficits in communication skills in poor households (Bando, Lopez Boo and Xi 2016). Otherwise, there is no evidence of sibling pairs fostering larger intra-household transfers of human capital.

One of the main conclusions that can be drawn from these results are that the overall

Table 4.8: Heterogeneous Bono 10000 impacts on Child Development(SD), by pairs of sibling gender

	ASQ	Communication	ASQ	Communication	ASQ	Communication
Panel A: Treatment interacted with Brothers						
Treatment * Brothers	-0.097	0.024	-0.134	0.047	-0.048	0.125
	[-0.4,0.2]	[-0.2,0.3]	[-0.4,0.1]	[-0.2,0.3]	[-0.3,0.2]	[-0.1,0.4]
Brothers	0.074	0.074	0.088	0.046	-0.010	-0.027
	[-0.2,0.3]	[-0.1,0.3]	[-0.1,0.3]	[-0.2,0.2]	[-0.3,0.2]	[-0.3,0.2]
Panel B: Treatment interacted with Sisters						
Treatment * Sisters	0.046	0.151	0.079	0.223	0.008	0.202
	[-0.2,0.3]	[-0.09,0.4]	[-0.2,0.3]	[-0.06,0.5]	[-0.3,0.3]	[-0.1,0.5]
Sisters	0.006	-0.070	-0.031	-0.146	-0.045	-0.181
	[-0.2,0.2]	[-0.3,0.2]	[-0.3,0.2]	[-0.4,0.1]	[-0.3,0.2]	[-0.5,0.1]
Panel C: Younger Sibling Male, Older Sibling Female						
Treatment *	0.107	0.247**	0.022	0.229*	0.102	0.198
Oldest Sibling is Female	[-0.1,0.3]	[0.010,0.5]	[-0.2,0.3]	[-0.04,0.5]	[-0.2,0.4]	[-0.1,0.5]
Older Sibling is Female	-0.079	-0.219**	-0.029	-0.195*	-0.020	-0.150
	[-0.3,0.1]	[-0.4,-0.02]	[-0.2,0.2]	[-0.4,0.02]	[-0.3,0.2]	[-0.4,0.1]
Panel D: Younger Sibling Female, Older Sibling Male						
Treatment *	0.088	0.121	0.096	0.121	0.099	0.147
Oldest Sibling is Male	[-0.1,0.3]	[-0.1,0.4]	[-0.1,0.3]	[-0.1,0.4]	[-0.2,0.4]	[-0.1,0.4]
Older Sibling is Male	-0.075	-0.069	-0.058	-0.030	-0.008	0.005
	[-0.3,0.1]	[-0.3,0.2]	[-0.3,0.2]	[-0.3,0.2]	[-0.2,0.2]	[-0.3,0.3]
HH Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects and lag ECD	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1122	1122	929	929	715	715

95% confidence intervals in brackets. Note: Standard errors clustered by aldea. All regressions include a constant.
The first sample includes all households in the panel. Sample 2 is all the households in the panel with at least one older sibling.
Sample 3 is the households with at least one eligible older sibling, and sample 4 is the households with 1-2 eligible siblings.
Model includes a selection of household and individual controls found in Table 4.1, as well as time,
baseline ECD measures, and department fixed effects.
Results of the fine motor skills, problem resolutions and social skills are available on request for brevity
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

treatment (household participation in Bono 10000) had a positive impact of about 0.17 SD on the language development of younger siblings compared with younger siblings of non-participants in non-treated aldeas. This result becomes larger and more significant across different specifications, suggesting that the impact is linked to the increased education/grade level of the older sibling, the window of opportunity in the age range for the younger sibling, and the gender of the older sibling, rather than simply arising as an income effect. The results arising from the different subsamples and the result from years of education and grade enrolled heterogeneity are consistent with the findings in Benedetti et al. (2016), where households with two eligible siblings show the greatest behavioural changes in terms of sending their children to school. If that is the case, we can say that the programme did create a positive spillover for the younger children and had additional impacts beyond the stated goals of the programme.

The fact that the communications domain is the one that is most impacted by the increased education of the older sibling is a logical conclusion. Older siblings have more school attainment and they have gained skills in interacting with their peers at a higher level of communication than before, and thus are better able to relate to others, including their siblings. Because of this, these children have high quality interactions with their younger siblings and their ability to communicate improves. Moreover, the effects seem to be mainly driven by the interactions of older females with younger boys, which are those that one would expect in a model of females being more involved in the parenting process and in which boys were the most disadvantaged at baseline.

We believe a crucial aspect of these evaluation is that siblings were not too far apart in age to be able to interact. The external validity of this study still needs to be tested in a setting where transfers are made to much older siblings.

4.7 Concluding remarks

The goal of this paper was to first explore whether a conditional cash transfer programme in Honduras had impacts on the cognitive development of children under 60 months, and then consider if the effect was amplified by intra-household spillovers spurred by the additional education attained by beneficiary children. We show that the programme led to significant ECD improvement in children, particularly in the communication dimension (even accounting for multiple hypotheses testing). These indirect effects are important to identify as they would lend additional support to the usefulness of cash transfers as effective means to reduce poverty. The stated goal of Bono 10000 was to improve school attendance for children between the ages of 6 to 18. This is an at-risk period of life for children in impoverished conditions, as the opportunity cost of going to school grows as children age and forego additional household income from labour market participation. Benedetti et al. (2016) showed that this main goal was achieved, with an overall treatment effect in school enrolment of 0.04 SD. In addition, there was a heterogeneous effect across different household sizes that is most significant in households with two eligible children at 0.052 SD.

After finding positive impacts of the larger transfer on communication skills, we examine in which households the programme works best. We show large effects (0.36 SD) of the programme when there are two eligible siblings in the household compared to the entire sample. When comparing across subsamples of the different compositions of eligible households, the effect is greatest (0.17 SD) in households with at least 1 child who is subject to the conditions of Bono 10000. This follows the previous results in Benedetti et al. (2016), indicating that smaller households benefit from the programme more than larger households. While Benedetti et al. (2016) argue that the fact that conditionalities were only binding for this group, it is also possible that this household size is the 'sweet spot', as the cash-transfer is enough to incentives households to send an additional child to school. As the older siblings are in school for longer, they learn more, and are better able to interact with their siblings, mainly in terms of their communication. Smith (1990) supports this conclusion, with the author's results indicating older siblings with larger vocabularies are better teachers for their younger siblings.

We then consider eligible child heterogeneity with different subsamples to determine if the spillovers are affected by gender, age difference, and the education level of the older child. These specifications add to the body of evidence supporting the presence of spillovers. The largest heterogeneity is found for those children with siblings completing 4-6 years of education. By keeping children in school and out of work, the cash transfer could be enabling stronger relationships between siblings and improving early childhood development outcomes. We also show that the effects are greatest in children who are between 25 and 36 months old, and where the oldest sibling is a female. More importantly, results seem to be driven by the interactions of older female siblings and younger boys, who are the most disadvantaged. These findings indicate a possible sensitive period of development and the distribution of responsibilities within the household which must be accounted for when considering the possibility of spillovers from a cash transfer. In addition, the CCT increased the consumption of healthy food and nutrients and maternal self-esteem, which could be responsible for some of the overall positive impact of Bono 10000 on early childhood development.

A caveat to the interpretation of positive intra-household spillovers is that the differences between the heterogeneous specifications and the standard treatment effects are statistically insignificant. The imprecision in the estimates is likely driven by small sample size. We conclude that a larger sample would provide more precise estimates which we believe should lead to stronger evidence of the effect of smarter children in the household.

The spillover effects measured in this evaluation perform reasonably well when compared to other CCTs in the region in terms of direct impacts stemming from policy. Evaluations of the direct effects of Bono de Desarollo Humano in Ecuador and Atención a Crisis in Nicaragua on young children led to estimated effect sizes of approximately 0.18SD (Paxson and Schady 2010; Macours, Schady, and Vakis 2012), which is comparable to the average we present here, but it is much larger if we look at some of the sub-samples. The estimates also demonstrate that the per household component of Bono 10000 provides an inflection point in comparison to the programme evaluated in Barrera-Osorio et al. (2008). In that study, the results indicated that a per-child randomised cash transfer led to some negative spillovers within the household. In ours, the per-household transfer seems to limit these effects, at least in terms of early childhood development, a fact supported by the positive labour results indicated in Benedetti et al. (2016).

An important result of any sort of cash transfer is that it is a cost-effective poverty reduction policy. Benedetti et al. (2016) performs a simple cost-effectiveness analysis for the main enrolment outcomes using the costs of PRAF-II as a guide.[36] They find that the full sample cost-effectiveness ratio is $13 per percentage point increase in

[36]Accurate costing of Bono 10000 is not available from the Honduran Government

enrolment. For the smaller households with the larger effects, this ratio falls to \$7 per percentage point increase in enrolment. Using this framework, the same analysis can be performed on the spillover results, showing a cost-effectiveness ratio of \$4, dropping to \$3 when conditioning on households with 1 to 2 eligible children. While this is a rough estimation of what the costs of the programme would be, it is possible to see that modest improvements in early childhood development can be made due to the spillover effects present in CCTs.

Altogether, our findings show that CCT's can be an important and cost effective way of improving early childhood development. The effects shown in this evaluation are important because they illustrate that there can be positive externalities in the household by increasing the human capital of one child, independent of any reallocation of resources or responsibilities. Our result adds to the evidence on the importance of conditions on the impacts and effectiveness of CCTs as a comprehensive tool for poverty reduction and the importance of considering intra-household spillover effects in any evaluation of policy.

5 Conclusions and future research

This thesis has studied early childhood development in Latin America in three specific ways. First, it examines the productivity of investments in the first five years of life with methods which are new to the field. Second, it advances the economic study of children's non-cognitive skill development deeper into the life cycle to the early adolescence. Lastly, it contributes to the intervention literature by evaluating a randomised control trial and identifying intra-household spillovers. These results offer policy relevant evidence of cash transfers being a useful tool for reducing poverty and improving early childhood development outcomes.

The first two studies use the Young Lives data from Peru as its primary data source. The first study asks when nutritional investment is most productive for skills development in the first five years of life, a period which is vital for children's success in the labour market (Almond 2006; Almond and Currie 2010). Disentangling the direction of causality between health and cognition is difficult without exogenous variation that is present with suitable instruments or experimental or quasi-experimental study designs. The challenge increases when considering two endogenous variables. To account for this identification concern, two instrumental variables methods are introduced. First, Anderson-Rubin (1951) Weak Identification Robust confidence regions help with inference in the standard case as strong identification in each of the endogenous variables is not guaranteed by a strong weak-identification statistic. The Lewbel (2012) method is then used to try and create exogeneity through instruments which are generated from assumptions about heteroskedasticity in the data. The results show that, at least in terms of associations, it is difficult to confidently determine if there are significant differences in the productivity of investments in the two periods. There is evidence of a significant indirect effect of early childhood health on later childhood skills development, suggesting that important neurocognitive features are developed early in life. The causal estimates are imprecisely estimated, and only weak inference can be achieved when combining the results together. This is especially important in relation to the existing literature, as it shows the importance of using robust methodologies to support inference made using instrumental variable methods. Lastly, using the Lewbel (2012) method in this setting highlights a path forward for longitudinal research which does not have access to suitable instruments or experimental or quasi-experimental variation.

The second study examines how relationships with parents and peers determine children's self-esteem and pride, while examining the role of socioeconomic status and locality in separating children in different categories. Self-esteem and pride are important concepts with respect to the non-cognitive skills that are necessary to be successful in the labour market. The descriptive analysis shows that socioeconomic gaps between upper and lower quintiles are prevalent in the early childhood, but begin to close around age 10. Locality gaps follow a similar trend but are of a smaller magnitude except for the parent-child relationship, where there is a large disparity between urban and rural households. With this information in mind, the paper continues by estimating the determinants of self-esteem and pride for the sample children. The results show no evidence of wealth gradients, but do show that relationships with parents and peers explain between 50 to 80 percent of the variance in children's self-esteem and pride. Analysis across the spectrum of self-esteem and pride shows that children of different levels of these psychosocial competencies react differently to inputs. The siblings difference model yields similar effect sizes to the initial estimations, meaning that the effect sizes are robust to the removal of household fixed effects. To conclude the analysis, the paper examines how much socioeconomic and locality gaps can be closed by equalising inputs between rich and poor and urban and rural children. Equalising the parent-child relationship leads to improvements of 21% in poor households and 80% for rural households at the mean, while distributional results are significant only in terms of self-esteem. The body of evidence shows that self-esteem and pride are related, but distinct constructs, which are reflections of a child's environment. The estimates suggest that the previous literature, which has not measured relationships, provides lower bound estimates of the effect of parents on their children's development. Policies which can improve these relationships, especially the parent-child relationship, are important for giving poorer children the essential human capital to overcome poverty in the future.

The final study evaluates a randomised cash transfer in Honduras on it spurring early childhood development through cash and intra-household education spillovers. The programme targeted children who were at risk to dropping out of education to switch to labour activities early in their academic careers. An important characteristic of the programme is that the cash transfer was per-household, rather than per-child. This changes the binding nature of the education condition, and caused effects which were different to previous studies in the field. Our results show that the programme itself had a significant impact of approximately 0.15 SD on early childhood development outcomes, specifically on children's communication abilities. The effect grows once the eligibility constraint is considered to 0.17 SD, which while larger than the treatment effect, is not significantly different to it. Examination of heterogeneous relationships shows that the development effects are largest when the beneficiary sibling has more education. More importantly, the effects seem to be mainly driven by the interactions of older female siblings with younger boys, which are those that one would expect in a model of females being more involved in the caregiving chores and in which boys present largest deficits in development. Both of these relationships increase the programme effect, strengthening the interpretation of positive intra-household spillovers as a result of the cash transfer. The results are comparable to direct early childhood development programme effects in Latin America, and support the efficacy of governments using cash transfers as means to reduce poverty and stimulate human capital growth in children.

While the three studies are distinct in specific topic matter, they all fall under the umbrella of early childhood development, and human capital development as a whole. The results offer some potential paths of future research. From the first chapter, the results emphasise the importance of using robust methodologies to ensure that inference is as accurate as possible. It also calls for additional research in experimental settings to help determine when nutritional investment is most important. When that setting is not feasible, a new methodology is introduced to overcome these challenges and offer new evidence. The second chapter extends the understanding of psychosocial competency development from the early childhood to the early adolescence. This provides a comprehensive picture of development, and suggests future research into the intra-household dynamics in play which mediates the gradients that are seen. Lastly, the third study evaluates a policy in place in Honduras. The identified spillover effects are important, because it offers evidence of a cash transfer having positive effects on human capital beyond the stated goal of a policy programme in a literature that mainly focuses on household allocation shifts. This is an important conclusion in its own right, and emphasises the importance of long-term studies to see if these effects persist onwards into the later childhood and adulthood.

References

Adermon, Adrien. 2013. Sibling Spillover in Education: Causal Estimates from a Natural Experiment. Doctoral dissertation, *Uppsala University*.

Almlund, Mathilde, Duckworth, Angela L., Heckman, James J. and Tim Kautz. 2011. Personality Psychology and Economics. *Handbook of the Economics of Education*, 4(1), 1-181.

Almond, Douglas. 2006. Is the 1918 Influenza pandemic over? Long-term effects of in utero Influenza exposure in the post-1940 US population. *Journal of Political Economy*, 114, 672-712.

Almond, Douglas and Janet Currie. 2010. Human Capital Development Before Age Five. *Handbook of Labor Economics*, 4(b), 1315-1486.

Almond, Douglas; Mazumder, Bhashkar and Reyn van Ewijk. 2015. In Utero Ramadan Exposure and Children's Academic Performance. *The Economic Journal*, 125(589), 1501-1533.

Anderson, Theodore W. and Herman Rubin. 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of Mathematical Statistics*, 20(1), 46-63.

Angelucci, Manuela and Vincenzo Di Maro. 2016. Programme evaluation and spillover effects. *Journal of Development Effectiveness*, 8(1), 22-43.

Anger, Silke. 2012. The intergenerational transmission of cognitive and non-cognitive skills during adolescence and young adulthood. *SOEP paper* no. 473.

Araujo, Maria Caridad; Bosch, Mariano and Norbert Schady. 2016. Can Cash Transfers Help Households Escape an Inter-Generational Poverty Trap? In *The Economics of Asset Accumulation and Poverty Traps.* University of Chicago Press.

Azubuike, Bridget and Kristine Briones. 2016. Young Lives Rounds 1 to 4 Constructed files. *Young Lives.*

Bando, Rosangela; López-Boo, Florencia and Xia Li. 2016. Sex-Differences in Language and Socio-emotional Skills in Early Childhood. *IDB Working Paper Series* No. IDB-WP-714.

Barham, Tara; Maccours, Karen and John A. Maluccio. 2013. Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions. *American Economic Review: Papers and Proceedings,* 103(3), 467-471.

Barker, David. J. P. 1995. Fetal origins of coronary heart disease. *British Medical Journal,* 311*,* 171.

Barrera-Osorio, Felipe; Bertrand, Marianne; Linden, Leigh L. and Francisco Perez-Calle. 2008. Conditional Cash Transfers in Education Design Features, Peer and Sibling Effects Evidence from a Randomized Experiment in Colombia. *The World Bank Policy Research Working Paper Series* No. 4580.

Baum, Christopher. F.; Schaffer, Mark E. and Steven Stillman. 2010. ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. *http://ideas. repec. org/c/boc/bocode/s425401. html.*

Baum, Christopher F.; Lewbel, Arthur, Schaffer, Mark E. and Oleksander Talavera. 2012. Instrumental variables estimation using heteroskedasticity-based instruments. *UK Stata Users Group Meetings*, 7, 13-14.

Behrman, Jere R. and Paul Taubman. 1986. Birth order, schooling, and earnings. *Journal of Labor Economics*, 4(3:2), S121-S145.

Behrman, Jere R. and Le Thuc Duc. 2014. Early-Childhood Growth Faltering, Post-Infancy Recovery and Educational Outcomes in Late Childhood: Evidence from Vietnam. *GCC Working Paper Series* No. GCC 14-04.

Benabou, Roland and Jean Tirole. 2003. Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70(3), 489-520.

Benedetti, Fiorella; Ibarrarán, Pablo, and Patrick McEwan. 2016. Do education and health conditions matter in a large cash transfer? Evidence from a Honduran experiment. *Economic Development and Cultural Change*, 64(4), 759-793.

Bhalotra, Sonia R. and Damien Clarke. 2016. The twin instrument. *IZA Discussion Paper* No. 10405.

Blinder, Alan S. 1973. Wage discrimination: reduced form and structural estimates. *The Journal of Human Resources*, 8(4), 436-455.

Boyden, Jo. 2016. Young Lives: an International Study of Childhood Poverty: Rounds 1-4 Constructed Files, 2002-2014. UK Data Service, 2nd Edition SN: 7483, *http://dx.doi.org/10.5255/UKDA-SN-7483-2.*

Branden, Nathaniel. 1994. Our urgent need for self-esteem: Our responses to events are shaped by whom and what we think we are. *Executive Excellence*, 11, 14.

Breusch, Trevor. S. and Adrian R. Pagan. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5), 1287-1294.

Campbell, Frances; Conti, Gabrielle; Heckman, James J.; Moon, Seong Hyeok; Pinto, Rodrigo; Pungello, Elizabeth and Yi Pan. 2014. Early Childhood Investments Substantially Boost Adult Health. *Science*, 343, 1478-1485.

Carneiro, Pedro, and Rita Ginja. 2014. Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start. *American Economic Journal: Economic Policy*, 6(4), 135-173.

Casale, Daniela and Chris Desmond. 2016. Recovery from stunting and cognitive outcomes in young children: evidence from the South African Birth to Twenty Cohort Study. *Journal of Developmental Origins of Health and Disease,* 7(2), 163-171.

Cobb-Clark, Deborah A. and Stefanie Schurer. 2012. The stability of big-five personality traits. *Economics Letters*, 115(1), 11-15.

Cook, Phillip J.; Dodge, Kenneth; Farkas, George; Fryer Jr, Roland. G.; Guryan, Jonathan; Ludwig, Jens; Mayer, Susan; Pollack, Harold and Laurence Steinberg. 2014. The (surprising) efficacy of academic and behavioral intervention with disadvantaged youth: results from a randomized experiment in Chicago. *National Bureau of Economic Research* No. w19862.

Cronbach, Lee J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.

Crookston, Benjamin; Dearden, Kirk; Alder, Stephen; Porucznik, Christina; Stanford, Joseph; Merrill, Ray; Dickerson, Ty and Mary Penny. 2011. Impact of early and concurrent stunting on cognition. *Maternal and Child Nutrition*, 7(4), 397-409.

Crookston, Benjamin. T.; Schott, Whitney; Cueto, Santiago; Dearden, Kirk. A.; Engle, Patrice; Georgiadis, Andreas; Lundeen, Elizabeth. A.; Penny, Mary E.; Stein, Areyh D. and Jere R. Behrman. 2013. Postinfancy growth, schooling, and cognitive achievement: Young Lives. *The American Journal of Clinical Nutrition*, 98(6), 1555-1563.

Cubel, Maria; Nuevo-Chiquero, Ana; Sanchez-Pages, Santiago and Marian Vidal-Fernandez. 2016. Do Personality Traits Affect Productivity? Evidence from the Laboratory. *The Economic Journal*, 126(592), 654-681.

Cueto, Santiago and Juan Leon. 2012. Psychometric characteristics of cognitive development and achievement instruments in Round 3 of Young Lives. *Young Lives Technical Note* 25.

Cunha, Flavio and James Heckman. 2007. The Technology of Skill Formation. *American Economic Review Papers and Proceedings*, 97(2), 31-47.

Cunha, Flavio; Heckman, James and Susanne M. Schennach. 2010. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3), 883-931.

Dahl, Ronald E. 2004. Adolescent brain development: a period of vulnerabilities and opportunities. Keynote address. *Annals of the New York Academy of Sciences*, 1021, 1-22.

Daly, Michael; Harmon, Colm P. and Liam Delaney. 2009. Psychological and biological foundations of time preference. *Journal of the European Economic Association*, 7(2-3), 659-669.

Daly, Michael; Delaney, Liam; Egan, Mark and Roy F. Baumeister. 2015. Childhood self-control and unemployment throughout the life span: Evidence from two British cohort studies. *Psychological Science*, 26(6), 709-723.

Darolia, Rajeev and Bruce Wydick. 2011. The Economics of Parenting, Self-esteem and Academic Performance: Theory and a Test. *Economica*, 78(310), 215-239.

Davidson, Russell and John G. MacKinnon. 2010. Wild bootstrap tests for IV regression. *Journal of Business & Economic Statistics,* 28(1), 128-44.

Dercon, Stefan and Pramila Krishnan. 2009. Poverty and the psychosocial competencies of children: evidence from the Young Lives sample in four developing countries. *Children Youth and Environments*, 19(2), 138-163.

Dercon, Stefan and Alan Sánchez. 2013. Height in mid childhood and psychosocial competencies in late childhood: Evidence from four developing countries. *Economics & Human Biology*, 11(4), 426-432.

Dercon, Stefan and Catherine Porter. 2014. Live Aid Revisited: Long-Term Impacts of the 1984 Ethiopian Famine on Children. *Journal of the European Economic Association*, 12(4), 927-948.

Duc, Le Thuc. 2011. Height and Cognitive Achievement of Vietnamese Children. *World Development*, 39(12), 2211-2220.

Duncan, Greg; Kalil, Ariel; Mayer, Susan; Tepper, Robin and Monique Payne. 2005. The apple does not fall far from the tree. *Unequal chances: Family background and economic success.* 23-79.

Dunn, Judy. 1983. Sibling relationships in early childhood. *Child Development*, 54(4), 787-811.

Dunn, Lloyd M.; Padilla, Eligio R.; Lugo, Delia E. and Leota M. Dunn. 1986. Manual del examinador para el Test de Vocabulario en Imágenes Peabody. *American Guidance Service.*

Dunn, Lloyd and Leota M. Dunn. 1997. Examiner's Manual for the PPVT-III. Form IIIA and IIIB. *American Guidance Service.*

Emran, M. Shahe and Zhaoyang Hou. 2013. Access to markets and rural poverty: evidence from household consumption in China. *Review of Economics and Statistics*, 95(2), 682-697.

Escobal, Javier and Eva Flores. 2008. An Assessment of the Young Lives Sampling Approach in Peru. *Young Lives Technical Note* 3.

Favara, Marta and Alan Sanchez. 2017. Psychosocial competencies and risky behaviours in Peru. *IZA Journal of Labor & Development,* 6(3).

Fernald, Lia C.; Kariger, Patricia; Engle, Patrice and Abbie Raikes. 2009. Examining early child development in low-income countries. *Washington DC: The World Bank.*

Fernald, Lia C. and Melissa Hidrobo. 2011. Effect of Ecuador's cash transfer program (Bono de Desarrollo Humano) on child development in infants and toddlers: a randomized effectiveness trial. *Social Science & Medicine*, 72(9), 1437-1446.

Fernald, Lia; Kariger, Patricia; Hidrobo, Melissa and Paul Gertler. 2012. Socioeconomic gradients in child development in very young children: Evidence from India, Indonesia, Peru, and Senegal. *Proceedings of the National Academy of Sciences*, 109(2), 17273-17280.

Ferreira, Francisco H.; Filmer, Deon and Norbert Schady. 2009. Own and sibling effects of conditional cash transfer programs: Theory and evidence from Cambodia. *World Bank Policy Research Paper* No. 5001.

Finlay, Keith; Magnusson, Leandro and Mark E. Schaffer. 2015. WEAKIV: Stata module to perform weak-instrument-robust tests and confidence intervals for instrumental-variable (IV) estimation of linear, probit and tobit models. *https://ideas.repec.org/c/boc/bocode/s457684.html.*

Firpo, Sergio.; Fortin, Nicole. M. and Thomas Lemieux. 2009. Unconditional quantile regressions. *Econometrica*, 77(3), 953-973.

Fiszbein, Ariel and Norbert Schady. 2009. *Conditional cash transfers: reducing present and future poverty.* World Bank Publications.

Fletcher, Jason; Hair, Nicole L. and Barbara L. Wolfe. 2012. Am I my brother's keeper? Sibling spillover effects: The case of developmental disabilities and externalizing behavior. *National Bureau of Economic Research* No. w18279.

Frisch, Ragnar and Frederick V. Waugh. 1933. Partial time regressions as compared with individual trends. *Econometrica*, 1(4), 387-401.

Galasso, Emanuela; Weber, Ann and Lia Fernald. 2017. Dynamics of child development: Analysis of a longitudinal cohort in a very low income country. *The World Bank Economic Review*, 1-26.

Galiani, Sebastian and Patrick McEwan. 2013. The heterogeneous impact of conditional cash transfers. *Journal of Public Economics*, 103, 85-96.

Georgiadis, Andreas and Priscila Hermida. 2014. Family Socio-economic Status, Mother's Psychosocial Skills, and Children's Human Capital: Evidence from Four Low-and Middle-income Countries. *Growing Up in Poverty*, Springer, 51-69.

Georgiadis, Andreas; Benny, Liza; Crookston, Benjamin. T.; Duc, L. T.; Hermida, Priscilla; Mani, Subha; Woldehanna, Tassew; Stein, Ayreh D. and Jere R. Behrman. 2017. Growth trajectories from conception through middle childhood and cognitive achievement at age 8 years: Evidence from four low-and middle-income countries. *Social Science & Medicine*, 2, 81-90.

Georgiadis, Andreas. 2017. Parental Background and Child Human Capital Development Throughout Childhood and Adolescence: Evidence from Four Low and Middle-Income Countries. *Young Lives Working Paper* 164.

Gilligan, Daniel; Margolies, Amy; Quiñones, Esteban and Shalini Roy. 2013. Impact evaluation of cash and food transfers at early childhood development centers in Karamoja, Uganda. *Washington, DC*: *IFPRI.*

Glewwe, Paul and Hanan Jacoby. 1995. An economic analysis of delayed primary school enrollment in a low income country: the role of early childhood nutrition. *The Review of Economics and Statistics*, 77(1), 156-169.

Glewwe, Paul and Elizabeth M. King. 2001. The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter? *The World Bank Economic Review*, 15(1), 81-113.

Gobierno De Uruguay. 2015. Encuesta de Nutrición, Desarollo Infantil, y Salud (ENDIS) 2015. *http://www.ine.gub.uy/encuesta-nacional-de-infancia-salud-nutricion-y-desarrollo-infantil-endis-2013.*

Grantham-McGregor, Sally; Cheung, Yin. B.; Cueto, Santiago; Glewwe, Paul; Richter, Linda; Strupp, Barbara and the International Child Development Steering Group. 2007. Developmental potential in the first 5 years for children in developing countries. *The Lancet*, 369(9555), 60-70.

Hahn, Jinyong and Jerry A. Hausman. 2003. IV estimation with valid and invalid instruments. *Annales d'Economie et de Statistique*, 79/80, 25-57.

Haney, Penny and Joseph A. Durlak. 1998. Changing self-esteem in children and adolescents: A meta-analytical review. *Journal of Clinical Child Psychology*, 27(4), 423-433.

Harkness, Sara and Charles M. Super. 1977. Why African children are so hard to test. *Annals of the New York Academy of Sciences*, 285(1), 326-331.

Heckman, James. and Yona Rubinstein. 2001. The importance of noncognitive skills: Lessons from the GED testing program. *The American Economic Review*, 91(2), 145-149.

Heckman, James J.; Stixrud, Jora and Sergio Urzua. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411-482.

Heckman, James J.; Moon, Seong Hyeok; Pinto, Ricardo; Savelyev, Peter A. and Adam Yavitz. 2010. The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, 94(1), 114-128.

Helmers, Christian and Manasa Patnam. 2011. The formation and evolution of childhood skill acquisition: Evidence from India. *Journal of Development Economics*, 95(2), 252-266.

Hoddinott, John and Bill Kinsey. 2001. Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics*, 63(4), 409-436.

Hoddinott, John; Alderman, Harold; Behrman, Jere R.; Haddad, Lawerence and Susan Horton. 2013. The economic rationale for investing in stunting reduction. *Maternal & Child Nutrition*, 9(S2), 69-82.

Janson, Harald and Jane Squires. 2004. Parent-completed developmental screening in a Norwegian population sample: a comparison with US normative data. *Acta paediatrica*, 93(11), 1525-1529.

Jann, Ben. 2017. ESTOUT: Stata module to make regression tables. Statistical Software Components. *https://ideas.repec.org/c/boc/bocode/s439301.html.*

Jürges, Hendrik. 2012. Collateral damage: The German food crisis, educational attainment and labor market outcomes of German post-war cohorts. *Journal of Health Economics*, 32(1), 286-303.

Kautz, Tim; Heckman, James. J.; Diris, Ron; Ter Weel, Bas and Lex Borghans. 2014. Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. *OECD Education Working Papers* No. 110.

Kessler, Daniel. 1991. Birth order, family size, and achievement: Family structure and wage determination. *Journal of Labor Economics*, 9(4), 413-426.

Kesternich, Iris; Siflinger, Bettina; Smith, James P. and Joachim K. Winter. 2014. The Effects of World War II on Economic and Health Outcomes across Europe. *Review of Economic Statistics*, 96(1), 103-118.

Kleibergen, Frank. 2005. Testing parameters in GMM without assuming that they are identified. *Econometrica*, 73(4), 1103-1123.

Kleibergen, Frank and Richard Paap. 2006. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1), 97-126.

Knudsen, Eric. 2004. Sensitive periods in the development of the brain and behaviour. *Journal of Cognitive Neuroscience*, 16(8), 1412-1425.

Krishnan, Priscila and Sofia Krutikova. 2013. Non-cognitive skill formation in poor neighbourhoods of urban India. *Labour Economics*, 24, 68-85.

Lagarde, Mylene; Haines, Andy and Natasha Palmer. 2009. The impact of conditional cash transfers on health outcomes and use of health services in low and middle income countries. *The Cochrane Library.*

Lea, Stephen E. and Paul Webley. 1997. Pride in economic psychology. *Journal of Economic Psychology*, 18(2), 323-340.

Lee, Chulhee. 2014. In utero exposure to the Korean War and its long-term effects on socioeconomic and health outcomes. *Journal of Health Economics*, 33, 76-93.

Lee, David S.. 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3), 1071-1102.

Lewbel, Arthur. 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, 30(1), 67-80.

Lincove, Jane A. and Adam Parker. 2016. The influence of conditional cash transfers on eligible children and their siblings. *Education Economics*, 24(4), 352-373.

Lindert, Peter H.. 1977. Sibling position and achievement. *The Journal of Human Resources*, 12(2), 198-219.

Loehlin, John. 2005. Resemblance in personality and attitudes between parents and their children. *Unequal chances: Family background and economic success*. 192-207.

Lovell, Michael C. 1963. Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, 58(304), 993-1010.

Lopez-Boo, Florencia. 2009. The Production Function of Cognitive Skills: Nutrition, Parental Inputs and Caste Test Gaps in India. *Young Lives Working Paper* 55.

—. 2013. Intercontinental Evidence on Socioeconomic Status And Early Childhood. *IDB Working Paper Series* No. IDB-WP-435.

—. 2016. Socio-economic status and early childhood cognitive skills: A mediation analysis using the Young Lives panel. *International Journal of Behavioral Development*, 40(6), 500-508.

Lopez Boo, Florencia and Maria E. Canon. 2014. Cognitive skills gaps in India: can (late) nutrition ameliorate them? *FRB of St. Louis Working Paper* No 2012-051C.

Macours, Karen; Schady, Norbert, and Renos Vakis. 2012. Cash transfers, behavioral changes, and cognitive development in early childhood: evidence from a randomized experiment. *American Economic Journal: Applied Economics*, 4(2), 247-273.

Majid, Muhammad F.. 2015. The persistent effects of in utero nutrition shocks over the life cycle: evidence from Ramadan fasting. *Journal of Development Economics*, 117, 48-57.

Martinez, Sebastian; Naudeau, Sophie and Vitor Pereira. 2012. The promise of preschool in Africa: A randomized impact evaluation of early childhood development in rural Mozambique. *Washington, DC: The World Bank.*

Millimet, Daniel L. and Jayjit Roy. 2016. Empirical tests of the pollution haven hypothesis when environmental regulation is endogenous. *Journal of Applied Econometrics*, 31(4), 652-677.

Mruk, Christopher. 2009. Defining self-esteem: An often overlooked issue with crucial implications. *Kernis, MH Self-esteem issues and answers: a sourcebook of current perspectives.* New York: Taylor & Francis, 10-17.

Nicoletti, Cheti and Birgitta Rabe. 2013. Inequality in Pupils' Test Scores: How Much do Family, Sibling Type and Neighbourhood Matter? *Economica*, 80(318), 197-218.

Nicoletti, Cheti and Birgitta Rabe. 2014. Sibling spillover effects in school achievement. *IZA Discussion Paper* No. 8615.

Nielsen, Helena and Juanna Joensen. 2015. Peer Effects in Math and Science. In *2015 Meeting Papers* No. 1343, Society for Economic Dynamics.

Oaxaca, Ronald. 1973. Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693-709.

Oreopoulos, Phillip; Brown, Robert S. and Adam M. Lavecchia. 2017. Pathways to education: An integrated approach to helping at-risk high school students. *Journal of Political Economy*, 125(4), 947-984.

Outes-Leon, Iñigo and Stefan Dercon. 2008. Survey attrition and attrition bias in Young Lives. *Young Lives Technical Note* 5.

Paxson, Christina and Norbert Schady. 2010. Does money matter? The effects of cash transfers on child development in rural Ecuador. *Economic Development and Cultural Change*, 59(1), 187-229.

Popli, Gurleen; Gladwell, Daniel and Aki Tsuchiya. 2013. Estimating the critical and sensitive periods of investment in early childhood: A methodological note. *Social Science & Medicine*, 97, 316-324.

Qureshi, Javaeria A. 2017. Additional returns to investing in girls' education: Impact on younger sibling human capital. *Accepted at Economic Journal.*

Rao, C. Radhakrishna. 1948. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proc. Cambridge Philos. Soc.,* 44, 50-57.

Robins, Richard W. and Kali H. Trzesniewski. 2005. Self-esteem development across the lifespan. *Current directions in Psychological Science*, 14(3), 158-162.

Romano, Joseph P. and Martin Wolf. 2005. Stepwise Multiple Testing as Formalized Data Snooping. *Econometrica*, 73(4), 1237-1282.

Rosenberg, Morris. 1965. Society and the adolescent self-image. *Princeton, NJ: Princeton University Press*, 11.

Rosenberg, M. and Simmons R. Black. 1971. The urban school child. *Washington, DC: American Sociological Association*.

Rubio-Codina, Marta; Attanasio, Orazio and Sally Grantham-McGregor. 2016. Mediating pathways in the socio-economic gradient of child development Evidence from children 6-42 months in Bogota. *International Journal of Behavioral Development*, 40(6), 483-491.

Sabia, Joseph J. 2007. The effect of body weight on adolescent academic performance. *Southern Economic Journal*, 73(4), 871-900.

Sanchez, Alan. 2017. The structural relationship between nutrition, cognitive and non-cognitive skills: evidence from four developing countries. *Economics & Human Biology*, 27(A), 33-54.

Schady, Norbert; Behrman, Jere; Araujo, Maria Caridad.; Azuero, Rodrigo; Bernal, Raquel; Bravo, David; Lopez-Boo, Florencia; Macours, Karen; Marshall, Daniela; Paxson, Christina and Renos Vakis. 2015. Wealth gradients in early childhood cognitive development in five Latin American countries. *Journal of Human Resources*, 50, 446-463.

Schaffer, Mark E.. 2010. XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. Statistical Software Components, *https://ideas.repec.org/c/boc/bocode/s456501.html*.

Schaffer, Mark E.. 2013. IVHETTEST: Stata module to perform Pagan-Hall and related heteroskedasticity tests after IV. Statistical Software Components, *https://ideas.repec.org/c/boc/bocode/s428801.html.*

Scholte, Robert S.; Van den Berg, Gerard J. and Marten Lindeboom. 2015. Long-run effects of gestation during the Dutch hunger winter famine on labor market and hospitalization outcomes. *Journal of Health Economics*, 39, 17-30.

Schonhaut, Luisa; Salinas, Patricia; Armijo, Iván; Schönstedt, Marianne; Álvarez, Jorge and Marcos Manríquez. 2009. Validación de un cuestionario autoadministrado para la evaluación del desarrollo psicomotor. *Revista chilena de pediatría*, 80(6), 513-519.

Smith, Thomas E.. 1990. Academic achievement and teaching younger siblings. *Social Psychology Quarterly*, 53(4), 352-363.

Squires, Jane and Diane Bricker. 1999. ASQ User´s guide. 2nd Edition. *Baltimore: Brookes Publishing Co.*

Squires, Jane, Heo, Kay H. and Paul Yovanoff. 2008. Cross-cultural adaptation of a pre-school screening instrument: comparison of Korean and US populations. *Journal of Intellectual Disability Research*, 52(3), 195-206.

Staiger, Douglas and James H. Stock. 1997. Instrumental Variables Regression with Weak Instruments. *Econometrica*, 65(3), 557-586.

Tierney, Joseph. P, Baldwin-Grossman, Jean and Nancy L. Resch. 1995. Making a Difference. An Impact Study of Big Brothers/Big Sisters. *Public/Private Ventures.*

Todd, Petra E. and Kenneth I. Wolpin. 2003. On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113, F3-F33.

— .2007. The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital*, 1(1), 91-136.

Tu, Yu-Kang; Tilling, Kate; Sterne, Jonathan A. and Mark S. Gilthorpe. 2013. A critical evaluation of statistical approaches to examining the role of growth trajectories in the developmental origins of health and disease. *International Journal of Epidemiology*, 42(5), 1327-1339.

Trzesniewski, Kali. H.; Donnellan, M. Brent. and Richard W. Robins. 2003. Stability of self-esteem across the life span. *Journal of Personality and Social Psychology*, 84(1), 205.

Van den Berg, Gerard; Lundborg, Petter; Nystedt, Paul and Dan-Olof Rooth. 2014. Critical Periods During Childhood and Adolescence: A Study of Adult Height Among Immigrant Siblings. *Journal of the European Economics Association*, 12(6), 1521-1557.

Walker, Susan P.; Wachs, Theodore D.; Grantham-McGregor, Sally; Black, Maureen M.; Nelson, Charles A.; Huffman, Sandra L.; Baker-Henningham, Helen; Chang, Susan M.; Hamadani, Jena D.; Lozoff, Betsy; Gardner, Julie M.M.; Powell, Christine A.; Rahmann, Atif and Linda Richter. 2011. Inequality in early childhood: risk and protective factors for early child development. *The Lancet*, 378, 1325-1338.

Woldehanna, Tassew; Behrman, Jere R. and Mesele W. Araya. 2017. The effect of early childhood stunting on children's cognitive achievements: Evidence from Young Lives Ethiopia. *Ethiopia Journal of Health Development*, 31(2), 75-84.

World Bank. 2013. *http://data.worldbank.org/indicator/SI.POV.GINI?locations=PE*

World Health Organization; UNICEF and others. 2009. WHO child growth standards and the identification of severe acute malnutrition in infants and children: joint statement by the World Health Organization and the United Nations Children's Fund Geneva. *World Health Organization*.

Young Lives. 2011. Young Lives Methods Guide: Sampling.

A Chapter 2 Appendix

Table A.1: Regressions of Nutrition on Cognitive Skill at Age 7-8 (non-imputed sample)

	OLS		IV	
Conditional Body Size, age 4-5	0.282***	0.0934***	0.162	-0.196
	[0.2,0.4]	[0.04,0.1]	[-0.1,0.5]	[-0.5,0.1]
Height for Age 6-18 months old	0.178***	0.0904***	0.250***	0.132**
	[0.1,0.2]	[0.05,0.1]	[0.10,0.4]	[0.009,0.3]
Parental Investment in Round 3		0.0713		0.0784*
		[-0.02,0.2]		[-0.01,0.2]
Mother's Height R1		-0.00527		
		[-0.01,0.004]		
Female		-0.0818*		-0.129**
		[-0.2,0.01]		[-0.3,-0.005]
Age of child (mths)		0.00593		0.0213*
		[-0.003,0.02]		[-0.004,0.05]
Household Size		0.00685		0.00188
		[-0.02,0.03]		[-0.03,0.03]
Caregiver has completed up to the secondary education		0.247***		0.281***
		[0.2,0.3]		[0.2,0.4]
Urban		0.158		0.175*
		[-0.04,0.4]		[-0.03,0.4]
Mother's age at birth		0.0144***		0.0199***
		[0.005,0.02]		[0.010,0.03]
Log Wealth of the Household		0.670***		0.746***
		[0.5,0.9]		[0.5,1.0]
Birth Order		-0.0908***		-0.111***
		[-0.1,-0.04]		[-0.2,-0.05]
Private School		0.235***		0.261***
		[0.1,0.4]		[0.1,0.4]
Observations	1315	1315	1315	1315
R-squared	0.118	0.317	0.0970	0.254
Under-ID			7.874	8.464
Weak-ID			9.093	10.17
J Test	0	0	0	0

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site.

Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different.

Children without measured birth weights removed

Column 1 contains no controls, Column 2 adds household and individual controls and regional fixed effects

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Regressions of Nutrition on Cognitive Skill at Age 7-8, using longer period of nutrition

	(1) OLS	(2)	(3)	(4) IV	(5)	(6)
Conditional Body Size, age 1-8	0.301***	0.0812***	0.0820***	0.143	-0.123	-0.126
	[0.2,0.4]	[0.02,0.1]	[0.02,0.1]	[-0.09,0.4]	[-0.3,0.09]	[-0.3,0.08]
Height for Age 1-2 years old	0.218***	0.0970***	0.0999***	0.269***	0.110*	0.113**
	[0.1,0.3]	[0.05,0.1]	[0.05,0.1]	[0.1,0.4]	[-0.0008,0.2]	[0.0009,0.2]
Parental Investment in Round 3		0.0721	0.0683		0.0652	0.0610
		[-0.02,0.2]	[-0.03,0.2]		[-0.03,0.2]	[-0.03,0.2]
Mother's Height R1		-0.00670	-0.00683			
		[-0.02,0.002]	[-0.02,0.002]			
Female		-0.115**	-0.116**		-0.134**	-0.135**
		[-0.2,-0.02]	[-0.2,-0.02]		[-0.2,-0.02]	[-0.2,-0.03]
Age of child (mths)		0.00594	0.00603		0.0148	0.0150
		[-0.005,0.02]	[-0.005,0.02]		[-0.004,0.03]	[-0.004,0.03]
Household Size		-0.000870	-0.000445		-0.00563	-0.00526
		[-0.02,0.02]	[-0.02,0.02]		[-0.03,0.02]	[-0.03,0.02]
Caregiver has completed up to the secondary education		0.262***	0.265***		0.275***	0.279***
		[0.2,0.3]	[0.2,0.4]		[0.2,0.4]	[0.2,0.4]
Urban		0.194*	0.206*		0.201**	0.214**
		[-0.005,0.4]	[-0.007,0.4]		[0.007,0.4]	[0.009,0.4]
Mother's age at birth		0.0149***	0.0147***		0.0184***	0.0183***
		[0.006,0.02]	[0.006,0.02]		[0.009,0.03]	[0.009,0.03]
Log Wealth of the Household		0.716***	0.720***		0.773***	0.779***
		[0.5,0.9]	[0.5,0.9]		[0.6,1.0]	[0.6,1.0]
Birth Order		-0.0996***	-0.0998***		-0.111***	-0.112***
		[-0.1,-0.05]	[-0.1,-0.05]		[-0.2,-0.06]	[-0.2,-0.06]
Private School		0.239***	0.244***		0.250***	0.256***
		[0.1,0.4]	[0.1,0.4]		[0.1,0.4]	[0.1,0.4]
Observations	1471	1471	1471	1471	1471	1471
R-squared	0.127	0.367	0.367	0.108	0.345	0.345
Under-ID				11.09	11.59	11.76
Weak-ID				16.48	18.02	18.92
J Test	0	0	0	0	0	0

95% confidence intervals in brackets. Robust Standard Errors, clustered on sentinel site. Wild bootstrap (1000 reps) hypothesis tests used to account for few clusters. Results are no different. Column 1 contains no controls, Column 2 adds household and individual controls and regional fixed effects. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

B Chapter 3 Appendix

Table B.1: Pooled Robustness

	Males	Females	Urban	Rural
Panel A: Self-Esteem				
Top Quintile of Wealth	-0.030	0.12*	0.061	0.25**
	[-0.2,0.1]	[-0.02,0.3]	[-0.06,0.2]	[0.05,0.4]
Maternal Pride	-0.051*	0.00084	-0.013	-0.038
	[-0.1,0.008]	[-0.05,0.06]	[-0.05,0.02]	[-0.1,0.04]
Parent Relationship Score	0.24***	0.40***	0.34***	0.29***
	[0.1,0.4]	[0.3,0.5]	[0.3,0.4]	[0.1,0.5]
Peer Relationship Score	0.42***	0.37***	0.41***	0.40***
	[0.3,0.5]	[0.3,0.4]	[0.3,0.5]	[0.3,0.5]
Caregiver's Education	0.0071	-0.0040	-0.0019	0.0090
	[-0.002,0.02]	[-0.02,0.010]	[-0.01,0.008]	[-0.01,0.03]
Panel B: Pride				
Top Quintile of Wealth	0.031	0.053	0.14**	0.012
	[-0.1,0.2]	[-0.1,0.2]	[0.006,0.3]	[-0.3,0.3]
Maternal Pride	-0.0019	0.021	0.043**	-0.035
	[-0.07,0.07]	[-0.02,0.06]	[0.0003,0.08]	[-0.1,0.03]
Parent Relationship Score	0.22***	0.44***	0.34***	0.30***
	[0.1,0.3]	[0.4,0.5]	[0.3,0.4]	[0.2,0.4]
Peer Relationship Score	0.30***	0.25***	0.26***	0.28***
	[0.2,0.4]	[0.2,0.4]	[0.2,0.3]	[0.1,0.4]
Caregiver's Education	0.0099	0.013***	0.013***	0.01
	[-0.002,0.02]	[0.004,0.02]	[0.003,0.02]	[-0.005,0.02]
Observations	481	525	636	370

95% confidence intervals in brackets. Robust standard errors clustered at community level.
Wild Bootstrap hypothesis testing confirms the results considering small cluster amount.
Controls include gender, household size, age in years, birth order, locality, standardized PPVT score, child's height, whether the house experienced a shock, and region fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

C Chapter 4 Appendix

Table C.1: Baseline differences in treatment and control groups

	(1) All children with ASQ scores Treatment/Control Difference	(2) At least 1 Older Sibling Treatment/Control Difference	(3) At Least 1 Eligible Older Sibling Treatment/Control Difference	(4) 1-2 Eligible Older Siblings Treatment/Control Difference
Age in Months	-0.01	0.10	0.53	0.79
	(0.68)	(0.69)	(0.81)	(1.03)
Total Income	267.07	327.12	540.77	1307.02
	(390.36)	(620.32)	(820.72)	(1177.58)
Household size	0.27*	0.15	0.14	-0.00
	(0.11)	(0.13)	(0.15)	(0.17)
Female	0.02	0.03	0.02	0.04
	(0.02)	(0.03)	(0.03)	(0.04)
Mother's education (years)	-0.00	-0.04	0.06	0.14
	(0.13)	(0.16)	(0.18)	(0.23)
Mother is literate	0.01	0.03	0.02	0.03
	(0.01)	(0.02)	(0.02)	(0.03)
Father's education (years)	-0.05	0.05	0.09	0.34
	(0.11)	(0.15)	(0.16)	(0.20)
Father is literate	0.00	0.01	0.03	0.03
	(0.02)	(0.02)	(0.03)	(0.03)
# of children age 0-5	0.03	-0.00	-0.03	-0.03
	(0.03)	(0.04)	(0.05)	(0.06)
# of children aged 6 to 18	0.08	0.02	0.06	0.03
	(0.06)	(0.08)	(0.08)	(0.08)
Dirt floor in dwelling (1/0)	-0.02	-0.03	-0.04	-0.07
	(0.02)	(0.03)	(0.03)	(0.04)
Piped water in dwelling (1/0)	0.03*	0.04*	0.05*	0.06
	(0.01)	(0.02)	(0.02)	(0.03)
Dwelling has bathroom or latrine (1/0)	-0.02	-0.02	-0.02	0.03
	(0.02)	(0.02)	(0.03)	(0.03)
Dwelling has electricity (1/0)	0.05*	0.03	-0.00	0.00
	(0.02)	(0.02)	(0.03)	(0.04)
Access to telephone (1/0)	-0.04**	-0.04	-0.04	0.01
	(0.02)	(0.02)	(0.02)	(0.03)
Rooms in Dwelling	-0.01	-0.01	0.05	0.03
	(0.06)	(0.07)	(0.08)	(0.11)
Adults in household who are Lenca	-0.00	0.01	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.02)
Dwelling only accessible by footpath (1/0)	-0.02	-0.02	-0.03	-0.02
	(0.02)	(0.02)	(0.03)	0.03
Observations	2520	1505	1122	715

Standard errors in Parentheses. Note: Column 1 includes all children who have calculable ASQ scores.

Column 2 includes all children who have an older sibling between the age of 6 to 18, regardless of eligibility.

Column 3 includes households who have at least one child between 6-18 years old who is enrolled lower than the 9th grade.

Column 4 includes the households where there is only one eligible older child. Differences calculated as the control group minus the treatment

Table C.2: Baseline ASQ scores by domain and age group

	Table C.2a: All HH									Table C.2b: HH with at least 1 older sibling								
	Raw			Internal st.		External st.				Raw			Internal st.		External st.			
Age Group	N	Mean	SD	Mean	SD	Mean	SD	% Delays	% Risk of delay	N	Mean	SD	Mean	SD	Mean	SD	% Delays	% Risk of delay
Communications																		
0-11 m.	511	51.60	11.38	0.00	1.00	0.27	1.07	4.31	9.00	383	52.10	11.21	0.04	0.97	0.31	1.06	4.18	8.88
12-23 m.	530	41.09	14.81	-0.01	0.99	-0.30	1.05	6.98	19.25	381	41.57	14.96	0.03	1.00	-0.26	1.06	7.09	18.64
24-35 m.	533	48.69	12.53	0.01	1.00	-0.21	1.04	8.82	11.07	405	48.52	12.50	0.00	0.99	-0.23	1.04	8.82	9.88
36-47 m.	464	52.85	9.49	0.01	0.98	0.13	0.85	2.80	7.76	366	52.93	9.65	0.02	1.00	0.14	0.87	2.80	6.83
48-60 m.	482	52.79	10.66	-0.01	1.02	-0.05	1.01	6.64	11.00	374	52.84	10.37	-0.01	1.00	-0.05	0.98	6.42	12.03
Average	2520	49.41	11.77	0.00	1.00	-0.03	1.00	5.91	11.61	1909	49.59	11.74	0.02	0.99	-0.02	1.00	5.86	11.25
Gross Motor Skills																		
0-11 m.	511	47.66	13.82	0.01	0.99	-0.37	1.38	14.29	11.75	383	48.11	13.86	0.03	0.99	-0.35	1.39	13.58	11.75
12-23 m.	530	50.65	13.22	-0.02	0.99	-0.34	1.32	9.06	11.32	381	50.86	13.79	0.00	1.03	-0.33	1.39	9.19	11.02
24-35 m.	533	50.95	12.63	0.00	1.01	-0.21	1.35	11.44	10.51	405	51.02	12.59	0.01	1.01	-0.21	1.34	11.36	10.12
36-47 m.	464	50.98	11.71	0.02	1.01	-0.30	1.27	10.34	11.64	366	50.85	11.64	0.00	1.01	-0.32	1.26	10.38	12.57
48-60 m.	482	53.18	11.25	0.04	0.97	-0.01	1.16	9.13	7.26	374	53.13	11.49	0.03	0.99	-0.02	1.18	10.16	6.68
Average	2520	50.69	12.53	0.01	1.00	-0.25	1.30	10.85	10.50	1909	50.80	12.67	0.01	1.01	-0.24	1.31	10.93	10.43
Problem Resolution																		
0-11 m.	511	49.49	15.30	-0.01	0.99	-0.16	-0.16	10.37	11.15	383	49.95	14.99	0.02	0.99	-0.12	1.36	9.14	11.23
12-23 m.	530	36.85	15.51	-0.02	0.99	-0.34	1.32	28.11	18.11	381	37.37	15.25	0.04	0.98	-1.09	1.49	28.11	17.32
24-35 m.	533	42.29	15.51	0.02	1.00	-0.71	1.39	18.95	17.45	405	42.15	15.28	0.02	0.98	-0.72	1.37	18.52	18.77
36-47 m.	464	39.95	17.01	-0.02	1.00	-1.08	1.52	27.80	17.24	366	40.34	17.11	0.00	1.01	-1.05	1.53	27.80	16.67
48-60 m.	482	33.77	16.89	0.03	0.99	-1.60	1.47	37.34	21.58	374	34.45	16.79	0.07	0.98	-1.54	1.46	35.03	22.19
Average	2520	40.47	16.05	0.00	0.99	-0.78	1.11	24.52	17.11	1909	40.85	15.88	0.03	0.99	-0.90	1.44	23.72	17.23
Social Skills																		
0-11 m.	511	50.75	10.99	-0.02	1.00	0.10	1.09	5.28	9.39	383	50.78	11.19	-0.02	1.02	0.09	1.13	6.01	8.09
12-23 m.	530	49.50	11.28	0.02	0.98	0.02	1.08	6.60	12.45	381	49.55	11.49	0.03	0.99	0.03	1.10	6.82	12.60
24-35 m.	533	45.60	12.56	0.02	1.00	-0.71	1.39	11.63	19.32	405	45.89	12.28	0.01	0.99	-0.43	1.21	10.37	20.49
36-47 m.	464	46.29	12.60	0.02	0.99	-0.49	1.21	13.58	17.89	366	46.27	12.60	0.01	0.99	-0.49	1.21	13.11	18.58
48-60 m.	482	48.29	12.00	0.07	0.99	-0.45	1.19	11.20	16.60	374	48.73	11.73	0.01	0.96	-0.40	1.16	10.70	15.51
Average	2520	48.09	11.89	0.02	0.99	-0.30	1.19	9.66	15.13	1909	48.24	11.86	0.01	0.99	-0.24	1.16	9.40	15.05

	Table C.2c: HH with at least one eligible older sibling									Table C.2d: HH with 1-2 eligible children								
	Raw			Internal		External				Raw			Internal		External			
Age Group	N	Mean	SD	Mean	SD	Mean	SD	% Delays	% Risk of Delay	N	Mean	SD	Mean	SD	Mean	SD	% Delays	% Risk of delay
Communications																		
0-11 M.	300	51.92	11.74	0.02	1.01	0.28	1.12	5.33	9.00	232	52.61	11.03	0.09	0.94	0.37	1.03	4.74	9.05
12-23 M.	300	41.48	15.10	0.04	1.01	-0.27	1.07	7.67	18.00	217	42.58	14.98	0.11	1.00	-0.21	1.06	6.45	18.43
24-35 M.	349	48.73	12.34	0.01	0.98	-0.21	1.03	9.17	8.60	267	49.21	12.29	0.05	0.97	-0.17	1.02	8.99	7.12
36-47 M.	316	53.03	9.69	0.03	1.00	0.15	0.87	3.16	6.33	245	53.87	8.85	0.11	0.91	0.22	0.78	1.22	6.53
48-60 M.	332	52.65	10.57	-0.03	1.03	-0.06	1.00	6.63	12.65	236	53.10	10.15	0.02	0.97	-0.02	0.96	5.93	11.44
Average	1597	49.56	11.88	0.02	1.01	-0.02	1.02	6.39	10.92	1197	50.4	12.26	0.08	0.96	0.04	1	5.47	10.51
Gross Motor Skills																		
0-11 M.	300	48.21	13.57	0.02	1.01	-0.35	1.42	13.00	11.33	232	48.85	13.55	0.08	0.99	-0.26	1.32	12.93	9.91
12-23 M.	300	50.48	14.17	-0.04	1.07	-0.39	1.45	10.33	11.33	217	50.78	13.64	-0.01	1.01	-0.36	1.38	10.60	11.06
24-35 M.	349	51.00	12.59	0.00	1.01	-0.20	1.32	11.46	10.03	267	51.65	12.21	0.06	0.97	-0.13	1.26	9.36	11.24
36-47 M.	316	50.88	11.61	0.01	1.01	-0.31	1.26	9.81	13.61	245	51.51	11.14	0.06	0.97	-0.24	1.21	8.57	13.06
48-60 M.	332	52.98	11.70	0.01	1.01	-0.03	1.20	10.24	7.23	236	53.60	11.51	0.07	1.00	0.03	1.19	9.75	5.93
Average	1597	50.71	12.73	0.00	1.02	-0.26	1.33	10.97	10.71	1197	51.28	12.41	0.05	0.99	-0.19	1.27	10.24	10.24
Problem Resolution																		
0-11 M.	300	49.74	15.75	0.00	1.02	-0.14	1.41	10.00	11.67	232	49.75	16.24	0.01	1.05	-0.12	1.46	10.34	10.78
12-23 M.	300	37.26	15.31	0.05	1.00	-1.10	1.49	27.67	16.67	217	38.42	14.92	0.11	0.97	-0.99	1.45	10.60	11.06
24-35 M.	349	45.78	15.30	0.01	0.98	-0.44	1.38	18.62	18.91	267	42.58	15.01	0.05	0.96	-0.68	1.35	17.23	18.73
36-47 M.	316	40.52	17.24	0.01	1.02	-1.03	1.54	9.81	13.61	245	41.54	17.16	0.08	1.02	-0.94	1.53	26.12	14.69
48-60 M.	332	34.92	16.98	0.11	0.99	-1.49	1.48	34.64	20.48	236	36.11	16.84	0.18	0.98	-1.39	1.47	30.93	21.19
Average	1597	41.64	16.12	0.04	1.00	-0.84	1.46	20.15	16.27	1197	41.68	16.73	0.08	0.99	-0.81	1.51	19.05	15.29
Social Skills																		
0-11 M.	300	50.41	11.81	-0.05	1.08	0.05	1.20	7.33	8.33	232	50.83	11.63	-0.01	1.07	0.11	1.17	6.90	6.47
12-23 M.	300	49.91	11.32	0.05	0.98	0.05	1.10	6.67	11.00	217	50.07	11.43	0.07	0.98	0.05	1.10	6.91	11.06
24-35 M.	349	45.78	12.52	0.01	1.01	-0.44	1.24	11.17	20.34	267	46.25	12.38	0.05	0.99	-0.39	1.22	10.49	19.85
36-47 M.	316	46.45	12.63	0.03	0.99	-0.47	1.22	27.22	16.46	245	46.71	12.64	0.05	0.99	-0.44	1.21	13.06	16.73
48-60 M.	332	48.93	11.67	0.13	0.95	-0.38	1.15	10.24	15.66	236	49.76	11.17	0.20	0.91	-0.30	1.10	8.90	13.98
Average	1597	48.30	11.99	0.03	1.00	-0.24	1.18	12.53	14.36	1197	48.72	11.85	0.06	0.97	-0.21	1.18	9.25	13.62

Table C.3: Lee Bounds for Table 3, Panels A and B

	Panel HH		Panel HH with at least 1 older sibling		Panel HH with 1-2 eligible older sibling	
	ASQ	Com.	ASQ	Com.	ASQ	Com.
Panel A: Treatment						
Lower Bound	-0.0053	0.084	-0.038	0.037	0.051	0.15
	[-0.12,0.11]	[-0.044,0.21]	[-0.20,0.12]	[-0.11,0.18]	[-0.12,0.22]	[-0.024,0.33]
Upper Bound	0.26***	0.29***	0.29**	0.34***	0.089	0.19
	[0.11,0.42]	[0.16,0.42]	[0.10,0.48]	[0.19,0.48]	[-0.067,0.24]	[-0.013,0.40]
Panel B: Treatment x Eligibility						
Lower Bound	0.099	0.13*	0.078	0.13*	0.051	0.15
	[-0.032,0.23]	[0.016,0.24]	[-0.093,0.25]	[0.0061,0.26]	[-0.092,0.19]	[-0.010,0.31]
Upper Bound	0.12	0.16*	0.22**	0.27***	0.089	0.19*
	[-0.050,0.29]	[0.022,0.30]	[0.063,0.37]	[0.12,0.43]	[-0.099,0.28]	[0.032,0.36]

95% confidence intervals in brackets. Columns follow subsamples in Table 4.3

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.4: Heterogeneity in Treatment by Household Size, all outcomes

	ASQ		Communication		Fine Motor skills		Problem Resolution		Social skills	
Panel A: Treatment status for number of children in the household										
Treatment * 1 child	0.104	0.062	0.024	0.032	-0.021	-0.065	0.106	0.073	0.152	0.141
	[-0.2,0.4]	[-0.2,0.3]	[-0.2,0.3]	[-0.2,0.3]	[-0.3,0.3]	[-0.3,0.2]	[-0.2,0.4]	[-0.2,0.3]	[-0.1,0.4]	[-0.1,0.4]
Treatment * 2 children	0.160	0.154	0.153	0.177*	-0.073	-0.060	0.181	0.154	0.154	0.124
	[-0.04,0.4]	[-0.04,0.3]	[-0.06,0.4]	[-0.02,0.4]	[-0.3,0.1]	[-0.2,0.1]	[-0.05,0.4]	[-0.07,0.4]	[-0.03,0.3]	[-0.06,0.3]
Treatment * 3 children	0.097	0.065	0.074	0.047	0.108	0.075	-0.007	-0.040	0.156	0.140
	[-0.2,0.4]	[-0.2,0.3]	[-0.2,0.3]	[-0.2,0.3]	[-0.1,0.3]	[-0.1,0.3]	[-0.2,0.2]	[-0.3,0.2]	[-0.09,0.4]	[-0.10,0.4]
Treatment * 4 children	0.075	0.016	0.223	0.190	0.031	-0.004	-0.002	-0.063	-0.057	-0.098
	[-0.3,0.4]	[-0.3,0.3]	[-0.07,0.5]	[-0.09,0.5]	[-0.3,0.3]	[-0.3,0.3]	[-0.3,0.3]	[-0.3,0.2]	[-0.4,0.3]	[-0.4,0.2]
Panel B: Treatment interacted with the number of eligible siblings in the household										
Treatment x 1 eligible child	0.020	0.013	0.113	0.120	-0.057	-0.067	-0.001	-0.022	0.024	-0.001
	[-0.2,0.2]	[-0.2,0.2]	[-0.1,0.3]	[-0.08,0.3]	[-0.3,0.2]	[-0.3,0.1]	[-0.2,0.2]	[-0.2,0.2]	[-0.2,0.2]	[-0.2,0.2]
Treatment x 2 eligible children	0.185	0.162	0.362***	0.360***	0.046	0.031	0.109	0.062	-0.033	-0.030
	[-0.06,0.4]	[-0.06,0.4]	[0.1,0.6]	[0.1,0.6]	[-0.2,0.3]	[-0.2,0.3]	[-0.1,0.3]	[-0.1,0.3]	[-0.3,0.2]	[-0.3,0.2]
Treatment x 3 eligible children	0.175	-0.002	0.250	0.143	-0.206	-0.297*	0.096	-0.059	0.317	0.201
	[-0.2,0.5]	[-0.3,0.3]	[-0.1,0.6]	[-0.2,0.5]	[-0.5,0.1]	[-0.6,0.01]	[-0.3,0.5]	[-0.4,0.3]	[-0.08,0.7]	[-0.1,0.5]
Treatment x 4 eligible children	-0.046	-0.154	-0.240	-0.248	0.034	-0.016	0.109	-0.002	-0.002	-0.105
	[-0.5,0.4]	[-0.5,0.2]	[-0.7,0.2]	[-0.6,0.1]	[-0.4,0.5]	[-0.4,0.4]	[-0.2,0.5]	[-0.3,0.3]	[-0.5,0.5]	[-0.5,0.3]
Observations	1505	1505	1505	1505	1505	1505	1505	1505	1505	1505

Note: Standard errors clustered by aldea. The first sample includes all households in the panel.

Sample 2 is all the households in the panel with at least one older sibling. Sample 3 is the households with at least one eligible older sibling,

All regressions include a constant. Model includes a selection of household and individual controls found in Table 4.1

as well as time, baseline ECD measures, and department fixed effects.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5: Bono 10000 impacts (SD), by child development domains (panel HH)

	ASQ		Communication		Motor Skills		Problem Resolution		Social Interactions	
Panel A: Treatment effects for each subsample										
Treatment	0.108	0.068	0.140**	0.134**	0.004	-0.020	0.067	0.019	0.093	0.065
	[-0.03,0.2]	[-0.05,0.2]	[0.005,0.3]	[0.01,0.3]	[-0.1,0.1]	[-0.1,0.1]	[-0.05,0.2]	[-0.08,0.1]	[-0.02,0.2]	[-0.04,0.2]
RW Hypothesis Rejection at 10%		Yes		Yes		Yes		Yes		Yes
Panel B: Treatment interacted with eligibility										
Treatment * Eligibility	0.085	0.036	0.179**	0.165**	-0.044	-0.075	0.044	-0.016	0.050	0.006
	[-0.08,0.3]	[-0.1,0.2]	[0.02,0.3]	[0.02,0.3]	[-0.2,0.1]	[-0.2,0.06]	[-0.1,0.2]	[-0.1,0.1]	[-0.10,0.2]	[-0.1,0.1]
RW Hypothesis Rejection at 10%		Yes		Yes		Yes		Yes		Yes
Panel C: Treatment interacted with higher and lower quintiles of income										
Treatment * poorest quintile	0.321**	0.105	0.317**	0.164	0.092	-0.008	0.220*	0.050	0.278**	0.131
	[0.06,0.6]	[-0.1,0.3]	[0.07,0.6]	[-0.05,0.4]	[-0.2,0.4]	[-0.2,0.2]	[-0.02,0.5]	[-0.2,0.3]	[0.02,0.5]	[-0.1,0.4]
Q1	-0.402***	-0.127	-0.283***	-0.099	-0.207**	-0.046	-0.313***	-0.105	-0.321***	-0.159*
	[-0.6,-0.2]	[-0.3,0.03]	[-0.5,-0.10]	[-0.3,0.05]	[-0.4,-0.01]	[-0.2,0.1]	[-0.5,-0.1]	[-0.3,0.05]	[-0.5,-0.1]	[-0.3,0.01]
Panel D: Treatment interacted with years of education										
Treatment * Years of Education	0.023	0.017	0.033**	0.032**	-0.005	-0.007	0.019	0.010	0.019	0.015
	[-0.0,0.05]	[-0.01,0.04]	[0.001,0.06]	[0.002,0.06]	[-0.04,0.03]	[-0.03,0.02]	[-0.008,0.05]	[-0.01,0.03]	[-0.01,0.05]	[-0.01,0.04]
Household Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time and Lagged ECD	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1505	1505	1505	1505	1505	1505	1505	1505	1505	1505

95% confidence intervals in brackets. Note: Cluster robust standard errors on aldea. All regressions include a constant.

The sample includes all households in the panel.

Model 2 includes a selection of household and individual controls found in Table 4.1 as well as time, baseline ECD measures and region fixed effects.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.6: Heterogeneity of Bono 10000 impact (SD), by child development domain (panel HH)

	ASQ		Communication		Motor Skills		Problem Resolution		Social Interactions	
Panel A: Treatment interacted with Terciles of Older Sibling's Education at Baseline										
Treatment * 1-3 Years of Education	-0.149	-0.149	-0.138	-0.138	0.085	0.085	-0.119	-0.119	-0.282**	-0.282**
	[-0.4,0.1]	[-0.4,0.1]	[-0.4,0.1]	[-0.4,0.1]	[-0.2,0.4]	[-0.2,0.4]	[-0.4,0.1]	[-0.4,0.1]	[-0.5,-0.02]	[-0.5,-0.02]
Treatment * 4-6 Years of Education	0.031	0.031	0.280**	0.280**	-0.132	-0.132	-0.058	-0.058	0.012	0.012
	[-0.2,0.2]	[-0.2,0.2]	[0.06,0.5]	[0.06,0.5]	[-0.3,0.07]	[-0.3,0.07]	[-0.3,0.1]	[-0.3,0.1]	[-0.2,0.2]	[-0.2,0.2]
Treatment * 7-8 Years of Education	0.142	0.142	0.173*	0.173*	0.008	0.008	0.109	0.109	0.111	0.111
	[-0.04,0.3]	[-0.04,0.3]	[-0.03,0.4]	[-0.03,0.4]	[-0.2,0.2]	[-0.2,0.2]	[-0.05,0.3]	[-0.05,0.3]	[-0.07,0.3]	[-0.07,0.3]
Panel B: Treatment interacted with Gender of older sibling (Female equal to 1)										
Treatment * Gender	0.109	0.073	0.178*	0.194**	0.022	-0.003	0.081	0.038	0.034	-0.013
	[-0.09,0.3]	[-0.09,0.2]	[-0.03,0.4]	[0.008,0.4]	[-0.2,0.2]	[-0.2,0.2]	[-0.09,0.3]	[-0.1,0.2]	[-0.1,0.2]	[-0.2,0.1]
Panel C: Treatment interacted with Age Difference of Older Sibling										
Treatment * Age Difference	0.009	0.006	0.013*	0.013*	0.001	-0.001	0.007	0.002	0.007	0.003
	[-0.005,0.02]	[-0.007,0.02]	[-0.002,0.03]	[-0.0006,0.03]	[-0.01,0.02]	[-0.01,0.01]	[-0.006,0.02]	[-0.009,0.01]	[-0.006,0.02]	[-0.008,0.01]
Observations	1122	1122	1122	1122	1122	1122	1122	1122	1122	1122

95% confidence intervals in brackets. Standard errors clustered by aldea.

The sample includes the households in the panel with at least one older sibling.

Model 2 includes a selection of household and individual controls found in Table 4.1 as well as time, baseline ECD measures, and department fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.7: Descriptive Statistics By Subsample and Panel for Additional Channels

	(1) All children with ASQ scores at baseline		(2) At least 1 Older Sibling		(3) At Least 1 Eligible Older Sibling		(4) 1-2 Eligible Older Siblings	
	mean	sd	mean	sd	mean	sd	mean	sd
Purchased Food Staples	0.83	0.37	0.84	0.37	0.84	0.36	0.85	0.36
Purchased Proteins	0.87	0.34	0.87	0.34	0.87	0.34	0.90	0.30
Purchased Fruits	0.44	0.50	0.44	0.50	0.43	0.50	0.46	0.50
Purchased Flour based product	0.98	0.14	0.98	0.14	0.98	0.13	0.98	0.12
Purchased Red Meat	0.36	0.48	0.38	0.48	0.37	0.48	0.39	0.49
Purchased Vegetables	0.83	0.37	0.85	0.36	0.85	0.36	0.86	0.34
Purchased Drinks	0.71	0.45	0.71	0.46	0.69	0.46	0.72	0.45
Purchased Tobacco, Liquor and Beer	0.08	0.27	0.08	0.27	0.07	0.26	0.08	0.27
Purchased Cooking Products	0.99	0.09	0.99	0.09	0.99	0.08	0.99	0.07
Spent Transfer on Food, Healthcare or Schooling	0.42	0.49	0.44	0.50	0.46	0.50	0.46	0.50
Has spent money on school products	0.14	0.34	0.14	0.35	0.15	0.35	0.15	0.36
You Hit Your Child	0.83	0.38	0.83	0.38	0.84	0.37	0.84	0.37
Your Spouse/Partner Hits Your Child	0.49	0.50	0.50	0.50	0.50	0.50	0.49	0.50
You saw your mother get hit when you were young	0.18	0.38	0.18	0.39	0.19	0.39	0.19	0.39
You were hit when you were young	0.28	0.45	0.29	0.45	0.28	0.45	0.28	0.45
Mother has decision making power	0.04	0.18	0.04	0.20	0.04	0.19	0.03	0.18
Mother's Labour Status	0.15	0.36	0.15	0.36	0.16	0.37	0.16	0.37
Observations	1505		1122		929		715	

Note: Column 1 includes all children in the sample. Column 2 includes all children who have an older sibling between the age of 6 to 18 and panel data, regardless of eligibility.
Column 3 includes households who have at least one child between 6-18 years old who is enrolled lower than the 9th grade and in the panel.
Column 4 includes households with one to two children who meet the constraint and are eligible and in the panel.
Column 5 includes the households where there is only one eligible older child and in the panel.

Table C.8: Descriptive Statistics By Subsample and Panel for Self-Esteem

	All children with ASQ scores at baseline	
	mean	sd
Positive Life Satisfaction	0.97	0.16
Thinks they are good enough	0.65	0.48
Thinks you have good qualities	0.95	0.21
Think you are able to do good	0.94	0.23
Are proud of yourself	0.58	0.49
I do not feel bad about myself	0.67	0.47
Think you have Value	0.91	0.29
Treated Respectfully	0.16	0.36
I do not think I am a failure	0.74	0.44
I have a positive attitude	0.94	0.23
Observations	652	

Note: Column 1 includes all children in the sample.
Column 2 includes all children who have an older sibling between the age of 6 to 18 and panel data, regardless of eligibility.
Column 3 includes households who have at least one child between 6-18 years old who is enrolled lower than the 9th grade and in the panel.
Column 4 includes households with one to two children who meet the constraint and are eligible and in the panel.
Column 5 includes the households where there is only one eligible older child and in the panel.

Table C.9: Significant differences between Treatment and Control in independent variables

	Treatment/Control Difference	Control Mean	Treatment Mean
Panel A: Expenditures			
Protein Purchases	0.05 **	0.84	0.89
Vegetable Purchases	0.09 ***	0.79	0.88
Bono Spent on Food, Health and Schooling	0.75 ***	0.03	0.78
Panel B: Maternal Self-Esteem			
Do you feel like a valued person	0.05 *	0.88	0.93
Do you feel respected	0.06 *	0.13	0.18
Do you have a positive attitude	0.04 *	0.92	0.96

A number of different variables including household spending on different food products,
schooling, household punishment on the child, and young mother's self-esteem were tested and were not significant.
The self-esteem measures have a sample size of 652 while the rest of the sample is 1505

Table C.10: Additional Channels: Spending

	ASQ		Communication		Fine Motor skills		Problem Resolution		Social skills	
Panel A: How did you spend the Bono 10000?										
Treatment *	-0.163	-0.100	0.091	0.074	-0.267	-0.074	-0.137	-0.089	-0.146	-0.212
spent Bono on Food, Healthcare, and Schooling	[-0.5,0.2]	[-0.4,0.2]	[-0.2,0.4]	[-0.3,0.4]	[-0.6,0.1]	[-0.4,0.3]	[-0.6,0.3]	[-0.5,0.3]	[-0.6,0.3]	[-0.7,0.3]
Panel B: Have you spent money on education?										
Treatment * spent money on education	0.016	0.022	0.171	0.169	-0.109	-0.095	-0.048	-0.021	0.008	-0.026
	[-0.3,0.3]	[-0.2,0.3]	[-0.1,0.5]	[-0.1,0.4]	[-0.4,0.2]	[-0.4,0.2]	[-0.3,0.2]	[-0.3,0.2]	[-0.3,0.3]	[-0.3,0.2]
Household Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time and Lagged ECD	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1505	1505	1505	1505	1505	1505	1505	1505	1505	1505

95% confidence intervals in brackets

Standard errors clustered by aldea. All regressions include a constant.

Sample is all households in the panel.

Model 2 includes a selection of household and individual controls found in Table 4.1 as well as time and baseline ECD measures.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.11: Mother's Labour Status

	ASQ		Communication		Fine Motor skills		Problem Resolution		Social skills	
Panel A: Mother's Labour Status										
Treatment *	-0.144	-0.168	-0.140	-0.134	-0.227	-0.220*	-0.107	-0.180	0.055	0.032
Mother is working	[-0.5,0.2]	[-0.4,0.08]	[-0.4,0.2]	[-0.4,0.1]	[-0.5,0.06]	[-0.5,0.04]	[-0.4,0.2]	[-0.4,0.05]	[-0.2,0.4]	[-0.2,0.3]
Household Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time and Lagged ECD	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1505	1505	1505	1505	1505	1505	1505	1505	1505	1505

95% confidence intervals in brackets

Note: Standard Errors clustered on aldea. All regressions include a constant.

Sample includes all households in the panel.

Model 2 includes a selection of household and individual controls found in Table 4.1 as well as time and baseline ECD measures.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table C.12: Self Esteem variables

	ASQ		Communication		Fine Motor skills		Problem Resolution		Social skills	
Panel A: Life Self-value										
Treatment * Values Self	0.059	0.079	0.088	0.125	-0.085	-0.080	0.082	0.099	0.033	0.039
	[-0.1,0.2]	[-0.09,0.2]	[-0.1,0.3]	[-0.05,0.3]	[-0.3,0.09]	[-0.3,0.10]	[-0.10,0.3]	[-0.07,0.3]	[-0.1,0.2]	[-0.1,0.2]
Panel B: Treated with respect										
Treatment * Respect	0.444^{*}	0.332^{*}	0.312	0.282	0.113	0.049	0.634^{***}	0.533^{***}	0.101	0.028
	[-0.02,0.9]	[-0.06,0.7]	[-0.1,0.7]	[-0.1,0.7]	[-0.3,0.5]	[-0.4,0.4]	[0.2,1.1]	[0.1,0.9]	[-0.4,0.6]	[-0.4,0.5]
Panel C: Do you have a positive attitude?										
Treatment * Positive Attitude	0.007	0.018	0.076	0.111	-0.133	-0.131	0.036	0.048	-0.012	-0.024
	[-0.2,0.2]	[-0.2,0.2]	[-0.1,0.3]	[-0.07,0.3]	[-0.3,0.04]	[-0.3,0.05]	[-0.1,0.2]	[-0.1,0.2]	[-0.2,0.2]	[-0.2,0.1]
Household Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Time and Lagged ECD	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	652	652	652	652	652	652	652	652	652	652

95% confidence intervals in brackets

Note: Standard Errors clustered on aldea. All regressions include a constant.

Sample includes all households in the panel.

Model 2 includes a selection of household and individual controls found in Table 4.1 as well as time and baseline ECD measures.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$